

2D akustische Wellenforminversion geführter Wellen
im Flachwasser

2D acoustic Full Waveform Inversion of guided waves in shallow water

Masterarbeit

für die Prüfung zum
Master of Science

am Karlsruher Institut für Technologie
des Studienganges Geophysik
Fakultät für Physik
Geophysikalisches Institut

von

Peter Michael Habelitz

Bearbeitungszeitraum	1. November 2016 - 31. Oktober 2017
Matrikelnummer	1733908
Erstgutachter	Prof. Dr. Thomas Bohlen
Zweitgutachter	Prof. Dr. Andreas Rietbrock

Erklärung der Eigenständigkeit

Erklärung

Gemäß §14 (5) der „Studien- und Prüfungsordnung des Karlsruher Instituts für Technologie (KIT) für den Masterstudiengang Geophysik“ vom 6. August 2015.

Ich versichere wahrheitsgemäß, die Arbeit selbstständig verfasst, alle benutzten Hilfsmittel vollständig und genau angegeben und alles kenntlich gemacht zu haben, was aus Arbeiten anderer unverändert oder mit Abänderungen entnommen wurde sowie die Satzung des KIT zur Sicherung guter wissenschaftlicher Praxis in der jeweils gültigen Fassung beachtet zu haben.

Ort

Datum

Unterschrift

Zusammenfassung

Wellenforminversion ist ein iteratives numerisches Optimierungsverfahren mit dem Ziel ein Modell der räumlichen Verteilung von Materialparameter zu finden, welches seismische Messdaten erklären kann. Dazu wird eine Zielfunktion definiert, welche den Abstand zwischen synthetisch modellierten Daten und observierten Daten quantifiziert. Mithilfe der Sensitivität dieser Zielfunktion auf Modelländerungen (Gradient) können Modellverbesserungen gefunden werden, die die Zielfunktion minimieren. Die Berechnung der synthetischen Daten erfordert die numerische Lösung der Wellengleichung, beispielsweise mit finiten Differenzen, was Rechen- und Speicherkapazitäten von heutigen Hochleistungsrechnern beansprucht.

Das Ziel dieser Masterarbeit ist die Anwendung der akustischen Wellenforminversion auf einen marinen Datensatz, der geführte Wellen der Wasserschicht enthält. Die Präprozessierung beschränkt sich auf die notwendige 3D-2D Transformation und Interpolation der Felddaten. Als v_P -Startmodell dient das Ergebnis einer Laufzeittomographie des Untersuchungsgebiets. Die Inversionsergebnisse können durch Vergleich mit einem Migrationsergebnis verifiziert werden. Langsamkeits-Frequenz-Spektren von synthetischen Seismogrammsektionen zeigen die erfolgreiche Rekonstruktion der geführten Wellen. Des Weiteren wurde die Präkonditionierung mit der inversen Diagonalen der approximierten Hessematrix zwischen ihrer Anwendung per Summe der Gradienten und per einzelner Schuss variiert. Es stellte sich heraus, dass die Anwendung per Schusssektion zu einer deutlichen Reduzierung von Quellartefakten führt.

Abstract

Full waveform inversion is an iterative numerical optimisation method with the aim of finding a model of the spatial distribution of material parameters, that can explain seismic measured data. Therefor an objective function is defined, which quantifies the distance between synthetic and observed data. With the sensitivity of this objective function to model changes (gradient), model improvements can be found that minimize the objective function. The calculation of synthetic data requires the numerical solution of the wave equation, for example, with finite differences, which demands computational and memory capacities of today's high-performance computers.

The aim of this master thesis is the application of acoustic full waveform inversion to a marine data set containing guided waves of the water layer. Preprocessing is limited to necessary 3D-2D transformation and interpolation of the field data. The result of a travel time tomography of the investigated area serves as a starting model. The inversion results can be verified by comparing them with a migrated model. Slowness-frequency spectra of synthetic seismogram sections show the successful reconstruction of guided waves. Furthermore the preconditioning with the inverse diagonal of the approximated Hessian was varied between their application per sum of gradients and per single shot. It turned out that the application to the gradient within each shot section leads to a clear reduction of source artifacts.

Inhaltsverzeichnis

Erklärung der Eigenständigkeit	I
Zusammenfassung	II
Abstract	III
Inhaltsverzeichnis	V
Abbildungsverzeichnis	VI
Tabellenverzeichnis	VIII
1. Einleitung	1
2. Theorie der Wellenausbreitung	2
2.1. Allgemeiner Spannungszustand	2
2.2. Impulserhaltung	4
2.3. Akustische Näherung	6
2.4. Dämpfung	6
2.5. Oberflächenwellen - Grenzflächenwellen	7
2.6. Slant-Stack	9
3. Finite Differenzen Methode	10
3.1. Diskretisierung von Raum und Zeit	10
3.2. Definition des <i>Standard Staggered Grid</i> (SSG)	11
3.3. Ablauf der Wellenfeldberechnung	13
3.4. Wahl der Diskretisierung	13
4. Wellenforminversion	18
4.1. Geophysikalische Perspektive	18
4.2. Mathematische Perspektive	19
4.3. Lineare und nicht lineare Inversionsprobleme	20
4.4. Inversion nicht linearer Probleme	22
4.5. Abstiegsrichtung	25
4.6. Bestimmung der Schrittweite - Liniensuche	40
4.7. Inversion des Quellsignals	44
4.8. Inversionsschema	47

5. Beschreibung der Felddaten	48
5.1. Akquisitionsgeometrie	48
5.2. Frequenzgehalt	50
5.3. Geführte Wellen	52
6. Präprozessierung	53
6.1. Qualitätskontrolle	53
6.2. 3D-2D Transformation	53
6.3. Interpolation der Felddaten	54
7. Wellenforminversion mit Referenzeinstellungen	56
7.1. Startmodell	57
7.2. Taper	57
7.3. Mehrstufige Inversionsstrategie	58
7.4. Beurteilungskriterien des Inversionserfolgs	58
7.5. Bewertung der Ergebnisse	60
8. Untersuchung von Quellsignalinversion und Präkonditionierung	65
8.1. Motivation	65
8.2. Beschreibung der Ergebnisse	66
9. Fazit	72
Literaturverzeichnis	73
A. Videos	76

Abbildungsverzeichnis

2.1. Darstellung des allgemeinen Spannungszustandes anhand des Volumenelements $\partial V = \partial x \partial y \partial z$	2
3.1. Darstellung des <i>Standard Staggered Grid</i> (SSG) mit der Verteilung von Material- und Wellenfeldparametern in 2 Dimensionen.	11
3.2. Gestalt der Gitterdispersion mit zunehmender Propagation (links -> rechts) bei Verwendung von 16 (oben), 4 (Mitte) und 2 (unten) Gitterpunkten pro minimaler Wellenlänge aus (Bohlen et al., 2015).	16
4.1. Visualisierung des nicht linearen Misfitverlaufs, sowie die Entwicklung bis zur 2. Ordnung im Punkt m_i	23
4.2. Veranschaulichung der Definition 4.16 für die Abstiegsrichtung \vec{p} . Der Misfit $S(\vec{m})$ ist als Projektion über Höhenlinien abhängig von einem zweidimensionalen Modell dargestellt.	25
4.3. Prinzipieller Verlauf des Misfits $F(\alpha_i)$, dargestellt in Abhängigkeit der Schrittweite α_i für eine fest gewählte Abstiegsrichtung \vec{p}_i	25
4.4. Vergleich zwischen Gradienten- und Newtonrichtung am Beispiel einer quadratischen (ellipsoidalen) Misfitfunktion in zwei Modellparameter m_1 und m_2 , eingezeichnet in die Projektionsfläche der Höhenlinien. Die Newton Richtung (grüner Pfeil) weist in Richtung des Minimums, während dies vom negativen Gradient (roter Pfeil) verfehlt wird.	38
4.5. Veranschaulichung der Armijo Regel, die ein hinreichender Abstieg sicherstellt. Der Bereich akzeptabler Schrittweiten ist blau gepunktet dargestellt.	42
4.6. Dargestellt ist der Misfitverlauf und dessen Ableitung, auf die sich die Krümmungsbedingung der Powell Wolfe Regel bezieht, um ineffiziente Schrittweiten zu verhindern. Das grüne Schrittweitenintervall genügt beiden Powell Wolfe Regeln.	42
4.7. Bestimmung einer effizienten Schrittweite $\alpha_{i,\min}$ mithilfe einer parabolischen Interpolation von drei Testschrittweiten $\alpha_{i,1}$, $\alpha_{i,2}$ und $\alpha_{i,3}$	44
4.8. Ablaufdiagramm der Wellenforminversion.	47

5.1.	Akquisitionsgeometrie der verwendeten Felddaten (relativ). In der Inversion wird die Projektion in die y - z Ebene verwendet. Für jede der 61 Quellen liegt eine Seismogrammsektion mit 239 Spuren vor. Die Nummerierung der Quellen erfolgt mit ansteigenden x -Werten.	49
5.2.	Normiertes Frequenzspektrum der Seismogrammsektion des 8. Schusses (vgl. Abbildung 5.3).	50
5.3.	Spurnormierte Seismogrammsektionen von Schuss 8, 19 und 31.	51
5.4.	Langsamkeits-Frequenz-Spektrum anhand der präprozessierten Felddaten von Schussektion 8.	52
6.1.	Effekt der 3D-2D Transformation auf Seismogramme der Schussektion 8 der Felddaten.	55
7.1.	v_P -Startmodell und das daraus abgeleitete Dichtemodell.	57
7.2.	v_P -Taper zur Unterdrückung von Modellupdates in der Wassersäule.	58
7.3.	Entwicklung des Misfit der Referenzlösung über dem Iterationsschritt (rot). Die senkrechten Linien kennzeichnen die Erhöhung der oberen Grenzfrequenz (4 Hz - 23 Hz) des Bandpassfilters, der vor jeder Iteration auf die observierten Daten angewendet wird.	60
7.4.	Vergleich zwischen observierten und synthetischen Seismogrammen des 8. Schusses für verschiedene Iterationen der Referenzlösung. Der zugehörige Misfit und die Frequenzstufen können in Abbildung 7.3 abgelesen werden.	61
7.5.	Langsamkeits-Frequenz-Spektrum anhand der synthetischen Seismogramme des finalen Modells (160. Iteration) der Referenzlösung von Schussektion 8 (links) und Komposition dieses Spektrums mit dem der präprozessierten Felddaten (vgl. Abbildung 5.4) (rechts). Das Spektrum der Felddaten ist grün dargestellt, das der synthetischen Daten violett. Bereiche mit ähnlicher Intensität sind als Graustufen gekennzeichnet.	62
7.6.	Überlagerung der invertierten Quellsignale aller 61 Schüsse von der Referenzlösung nach 160 Iterationen.	63
7.7.	Modelle zur Bewertung des Inversionsergebnisses der Referenzlösung. a) Startmodell und b) Inversionsergebnis nach 160 Iterationen der gesamten Ausdehnung. c) Migrationsmodell. d) Inversionsergebnis als Ausschnitt von b). e) und f) zeigen Merkmale, die sich sowohl im Inversions- als auch im Migrationsergebnis finden, wobei in f) beide überlagert sind.	64

8.1.	Entwicklung des Misfits aller vier Inversionen über dem Iterationsschritt (rot). Die senkrechten Linien markieren die Iterationen bei denen die obere Grenzfrequenz (4 Hz - 23 Hz) des Bandpassfilters erhöht wird. Dieser Filter wird vor jeder Iteration auf die observierten Daten angewendet.	67
8.2.	Vergleich zwischen observierten und synthetischen Seismogrammen des 8. Schusses. Die Nummer der Iteration wurde so gewählt, dass jede Iteration am Abbruchkriterium des Frequenzbandes mit Grenzfrequenz von 16 Hz angelangt ist. Der zugehörige Misfit und die Frequenzbänder können in Abbildung 8.1 abgelesen werden.	69
8.3.	Überlagerung der invertierten Quellsignale aller 61 Schüsse für jede der 4 Inversionen nach 160 Iterationen. Um ein Vergleich untereinander zu ermöglichen, wurde die Normierung aller 244 Signale mit einem einzigen Faktor vorgenommen.	70
8.4.	Invertierte v_P -Modelle nach 160 Iterationen.	71
A.1.	Evolution verschiedener Inversionsergebnisse der Referenzlösung (vgl. Kapitel 7): Misfit, v_P -Modell, Vergleich zwischen synthetischen und observierten Seismogrammen, Überlagerung von Migrationsergebnis und v_P -Modell, invertierte Quellsignale aller Schüsse und v_P -Gradient.	76
A.2.	Simulation der akustischen Wellenausbreitung im finalen Modell (160. Iteration, Einstellung: STFI pro Frequenzband, Präkonditionierung per Schuss) von Schuss 1. Zur besseren Sichtbarkeit der Ausbreitungswege sind die Amplituden normiert. Gut zu Erkennen sind Tauchwellen, welche Informationen von tieferen Schichten sammeln und die multiplen Reflexionen zwischen Oberfläche und Meeresboden, die zur Ausbildung der geführten dispersiven Wellen führen.	77

Tabellenverzeichnis

3.1.	Vergleich der analytisch und numerisch (Gl.: 3.5), mit verschiedenen Gitterabständen $\Delta h = \frac{\lambda}{n}$, berechneten Ableitung von $v_x = \sin\left(2\pi\frac{x}{\lambda}\right)$	15
3.2.	Anzahl n der benötigten Gitterpunkte pro minimaler Wellenlänge für verschiedene Operatorlängen und -typen, um die Gitterdispersion weitgehend zu unterdrücken.	15

3.3. Konstante γ des Courant-Friedrichs-Lewy-Kriteriums für die FD-Ordnungen 2 bis 12.	17
5.1. Übersicht zu den Akquisitionsparameter der Felddaten	48
7.1. Referenzeinstellungen der Wellenforminversion	56
7.2. Inversionsstufen der <i>multiscale</i> Strategie mit der sich progressiv erhöhen- den oberen Grenzfrequenz f_{high} , der unteren Grenzfrequenz f_{low} und dem Abbruchkriterium Prog_S	59

1. Einleitung

Wissenschaftliche Durchbrüche sind manchmal dem Zufall geschuldet. Ein bekanntes Beispiel ist die Entdeckung des Penicillins durch Alexander Fleming. Weniger populär, doch für die Geophysik von enormer Bedeutung, ist die erste Aufzeichnung eines Fernbebens durch den Astronomen Ernst von Rebeur-Paschwitz. Mit seinem Instrument wollte er ursprünglich die Bewegung von Himmelskörpern messen. Es war der Beginn der modernen Seismologie und ermöglichte in der Folge die Auswertung von Laufzeitinformationen, was zu Rückschlüssen auf den Erdaufbau befähigte ([Gutenberg, 1914](#); [Lehmann, 1936](#)). Technische Verbesserungen der Seismometer im 20. Jahrhundert erlaubten die Verwendung der aufgezeichneten Amplitudeninformation zur Erhöhung der Modellauflösung. Als in der zweiten Hälfte des 20. Jahrhunderts die Theorie zur Wellenforminversion aufgestellt wurde, fehlte es für deren raschen Verbreitung zunächst noch an verfügbarer Rechenleistung zur Modellierung von seismischen Daten. Diese werden beginnend von einem Startmodell durch lokale Optimierung an die observierten Daten angepasst. Begünstigt durch die kontinuierlichen Entwicklungen in der Computertechnologie konnten maßgebende Fortschritte erzielt werden. Die Wellenforminversion ist heute in der Lage hoch aufgelöste Modelle zu erzeugen.

Anmerkungen

Zur Durchführung der Inversion wurde die am Karlsruher Institut für Technologie entwickelte Software IFOS2D verwendet.

Diese Arbeit wurde an der vom Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg und der Deutschen Forschungsgemeinschaft (DFG) finanzierten Rechenressource ForHLR II durchgeführt.

Die Masterarbeit selbst wurde mit \LaTeX erstellt, die Vektorgrafiken mit TikZ. Zur Präprozessierung und Nachbearbeitung der Ergebnisse wurde MatLab verwendet. Im Anhang befinden sich eingebundene Videos, deren Abspielbarkeit mit Adobe Acrobat Reader DC Version 2017 unter MacOS getestet wurde. Zusätzlich müssen sich die Videos in der vorgegebenen Ordnerstruktur befinden.

2. Theorie der Wellenausbreitung

In diesem Kapitel werden die theoretischen Grundlagen für die physikalische Beschreibung der Wellenausbreitung in elastischen Medien gelegt. Die akustische Wellengleichung lässt sich hieraus als Vereinfachung ableiten. Die Bestandteile der Herleitung sind die allgemeine Spannungs-Dehnungs-Beziehung, auch Materialgesetz genannt, und die Impulserhaltung.

2.1. Allgemeiner Spannungszustand

Anders als nicht viskose Fluide können viskose Fluide und Festkörper unter Schubspannung stehen und die Normalspannung in den 3 Raumrichtungen sind im Allgemeinen verschieden. Dies ist durch die vorhandene inneren Reibung zu erklären.

Der Spannungszustand wird anhand eines infinitesimalen Volumenelements $\partial V = \partial x \partial y \partial z$ in Abbildung 2.1 dargestellt.

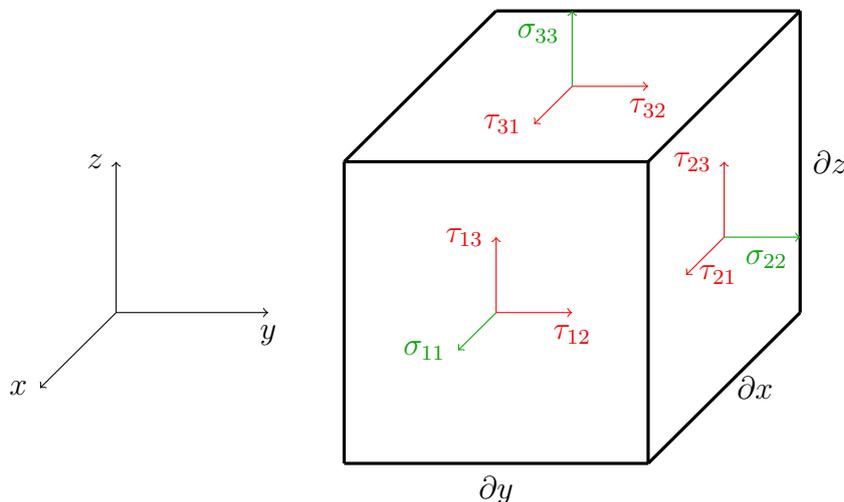


Abbildung 2.1.: Darstellung des allgemeinen Spannungszustandes anhand des Volumenelements $\partial V = \partial x \partial y \partial z$.

Die Indizes des Spannungstensors σ_{ij} bezeichnen die Richtung i der Flächennormalen, auf welche die Spannung wirkt und die Richtung j der Spannung selbst. σ_{ii} sind Normalspan-

nungen mit der Konvention $\sigma_{ii} > 0$: Zug, $\sigma_{ij} < 0$: Druck. σ_{ij} mit $i \neq j$ werden auch mit τ_{ij} bezeichnet und stellen die Schubspannungen dar.

$$\sigma_{ij} = \begin{pmatrix} \sigma_{11} & \tau_{12} & \tau_{13} \\ \tau_{21} & \sigma_{22} & \tau_{23} \\ \tau_{31} & \tau_{32} & \sigma_{33} \end{pmatrix} \quad (2.1)$$

Da der Spannungszustand Körper im statischen Gleichgewicht darstellt, ist für jede Spannung auf den drei verdeckten Seiten je eine entgegen gerichtete Gegenspannung vorhanden. Damit das Volumenelement nicht in Rotation gerät muss also gelten, dass $\sigma_{ij} = \sigma_{ji}$ (Drehimpulserhaltung). Daraus folgt, dass σ_{ij} symmetrisch ist und 6 unabhängige Spannungen besitzt. Durch das Hookesche Gesetz wird der Zusammenhang zwischen Spannung und Dehnung hergestellt. Für kleine Dehnungen gegenüber der Wellenlänge, was bei der Propagation von seismischen Wellen stets gegeben ist, hängt jede Spannungskomponente näherungsweise linear von den Dehnungen ab.

$$\sigma_{ij} = C_{ijkl}\epsilon_{kl} \quad (2.2)$$

σ_{ij} : Spannungstensor
 ϵ_{kl} : Verzerrungstensor
 C_{ijkl} : Elastizitätstensor

Der Elastizitätstensor C ist 4-dimensional und hat im Allgemeinen $3^4 = 81$ Elemente. Er besitzt die Symmetrien $C_{ijkl} = C_{jikl} = C_{ijlk} = C_{klij}$, wodurch sich die Zahl der unabhängigen Komponenten auf 21 reduziert. Bei isotropen Medien bleiben schließlich noch 2:

$$C_{ijkl} = \lambda\delta_{ij}\delta_{kl} + \mu(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}) \quad (2.3)$$

Somit lässt sich das generalisierte Hookesche Gesetz für linear elastische, isotrope Medien formulieren:

$$\sigma_{ij} = \lambda\epsilon_{kk}\delta_{ij} + 2\mu\epsilon_{ij} \quad (2.4)$$

$$\epsilon_{ij} = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \quad (2.5)$$

Hier bezeichnet λ die erste Lamé-Konstante und μ die zweite Lamé-Konstante (auch Schubmodul genannt). δ_{ij} ist das Kronecker-Delta. ϵ_{ij} bezeichnet nun den linearisierten Verzerrungstensor. u_i ist die Verschiebung eines Partikels in Richtung i .

2.2. Impulserhaltung

Neben der Spannungs-Dehnungs-Beziehung ist die Impulserhaltung (Kraftdichtebilanz) der Kern der elastischen Wellenausbreitung. Sie entspricht prinzipiell der differentiellen Formulierung des dynamischen Gleichgewichts pro Volumeneinheit. Man betrachte dazu ein infinitesimales Volumen mit der Kantenlänge ∂x_i . Auf ihn wirken zwei Arten von Kräften: Volumenkraft und Oberflächenkraft. (Landau and Lifschitz, 1997)

Die Volumenkraft, Trägheitskraft und Gravitationskraft, wirken auf den Körper aufgrund seines massebehafteten Volumens. Die Trägheitskraft lässt sich mit dem zweiten Newtonschen Axiom in seiner ursprünglichen Formulierung schreiben als:

$$\frac{dp}{dt} = \frac{d(\partial m \cdot v_i)}{dt} = \underbrace{\frac{d(\partial m)}{dt}}_{=0, \text{ da } m \text{ unabhängig von } t} v_i + \partial m \frac{dv_i}{dt} \quad (2.6)$$

und als Trägheitskraftdichte mit $v_i = v_i(x_k, t)$:

$$f_{T,i} = \frac{\partial m}{\partial V} \frac{dv_i}{dt} = \rho \frac{dv_i}{dt} = \rho \left[\frac{\partial v_i}{\partial t} + \underbrace{v_k \frac{\partial v_i}{\partial x_k}}_{\approx 0, \text{ da Verschiebungen klein gegenüber Wellenlängen}} \right] \approx \rho \frac{\partial v_i}{\partial t} \quad (2.7)$$

Für elastische Raumwellen kann die Gravitationskraft vernachlässigt werden. Sie spielt lediglich für sehr niederfrequente Eigenschwingungen der Erde eine Rolle. Die Oberflächenkraft wirkt auf den Körper aufgrund des Spannungszustandes. Die Berechnung erfolgt anhand der Kräftebilanz für alle drei Raumrichtungen. Zunächst wird die Oberflächenkraft als Integral des Spannungsvektors \vec{s} über die Oberfläche eines Volumenelements ∂V dargestellt.

$$\vec{F}_O = \int_A \vec{s} \, dA \quad (2.8)$$

Allgemein lässt sich der Spannungsvektor, der auf einer beliebigen Fläche steht, durch Multiplikation des transponierten Spannungstensors mit der Flächennormalen \vec{n} berechnen. Hier ist zu beachten, dass der Spannungstensor aufgrund der Drehimpulserhaltung symmetrisch ist (s. Kapitel 2.1) und somit gilt $\sigma^T = \sigma$

$$\vec{s} = \sigma^T \vec{n} \quad (2.9)$$

Mit dem gerichteten Flächenelement $d\vec{A} = \vec{n} \, dA$ lässt sich die Oberflächenkraft umschrei-

ben.

$$\vec{F}_O = \int_A \sigma \vec{n} \, dA = \int_A \sigma \, d\vec{A} \quad (2.10)$$

Der Integralsatz von Gauß $\int_A \vec{F} \, d\vec{A} = \int_V \operatorname{div} \vec{F} \, dV$ ermöglicht es das Oberflächenintegral in ein Volumenintegral zu überführen. Dies ist hier vorteilhaft, da so durch das Volumenelement geteilt werden kann, wodurch man die gewünschte Größe einer Oberflächenkraftdichte \vec{f}_O erhält.

$$\vec{F}_O = \int_A \sigma \, d\vec{A} = \int_V \operatorname{div} \sigma \, dV \quad (2.11)$$

$$\vec{f}_O = \frac{\vec{F}_O}{dV} = \operatorname{div} \sigma \quad \vec{f}_{O,i} = \frac{\partial \sigma_{ij}}{\partial x_j} \quad (2.12)$$

Schließlich kann das Kraftdichtegleichgewicht $f_{T,i} = f_{O,i}$ aus Gleichung 2.7 und 2.12 zur elastischen Bewegungsgleichung für linear elastische, isotrope Medien ausformuliert werden.

$$\rho \frac{\partial v_i}{\partial t} = \frac{\partial \sigma_{ji}}{\partial x_j} \quad (2.13)$$

Ableiten der Materialgesetze nach der Zeit führt zusammen mit Gl. 2.13 auf die sogenannte Spannungs-Geschwindigkeits Formulierung, welche sich für die Behandlung mit finiten Differenzen eignet:

$$\frac{\partial \sigma_{ij}}{\partial t} = \lambda \frac{\partial \epsilon_{rr}}{\partial t} \delta_{ij} + 2\mu \frac{\partial \epsilon_{ij}}{\partial t} \quad (2.14)$$

$$\frac{\partial \epsilon_{ij}}{\partial t} = \frac{1}{2} \left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right) \quad (2.15)$$

Die Bewegungsgleichung nach Cauchy stellt eine weitere Formulierung der Wellengleichung dar, die man durch Ausnutzung der Symmetrie $C_{ijkl} = C_{ijlk}$ erhält. Dazu wird Gl. 2.5 und Gl. 2.2 in Gl. 2.13 eingesetzt und $\frac{\partial v_i}{\partial t} = \frac{\partial^2 u_i}{\partial t^2}$ verwendet.

$$\begin{aligned} \rho \frac{\partial^2 u_i}{\partial t^2} &= \frac{1}{2} \frac{\partial}{\partial x_j} \left(C_{ijkl} \frac{\partial u_k}{\partial x_l} + C_{ijkl} \frac{\partial u_l}{\partial x_k} \right) \\ &= \frac{1}{2} \frac{\partial}{\partial x_j} \left(C_{ijkl} \frac{\partial u_k}{\partial x_l} + C_{ijlk} \frac{\partial u_k}{\partial x_l} \right) \\ &= \frac{\partial}{\partial x_j} \left(C_{ijkl} \frac{\partial u_k}{\partial x_l} \right) \end{aligned} \quad (2.16)$$

Im zweiten Summand werden die doppelten Indizes k und l umbenannt ($k \rightarrow l, l \rightarrow k$), nicht vertauscht. Mit $C_{ijkl} = C_{ijlk}$ folgt die Identität der Summanden.

2.3. Akustische Näherung

Die akustische Vereinfachung der Wellenausbreitung ist durch Medien ohne Scherwiderstand ($\mu = 0$) gegeben, wie sie nicht viskose Fluide darstellen. Aber auch Festkörper können akustisch genähert werden, womit Scherwellen vernachlässigt werden. Der Elastizitätstensor aus Gleichung 2.3 vereinfacht sich zu $C_{ijkl} = \lambda\delta_{ij}\delta_{kl}$. Mit dem Hookschen Gesetz (Gl. 2.2) reduziert sich der Spannungstensor σ_{ij} auf seine Diagonalelemente, die zudem alle den Wert $\epsilon_{kk} = \frac{\partial u_k}{\partial x_k}$ annehmen. Da der Druck p allgemein als Mittelwert der Normalspannungen σ_{ii} definiert ist ($p := \frac{1}{3}\sigma_{ii}$) gilt

$$\sigma_{ij} = \lambda\delta_{ij}\delta_{kl}\epsilon_{kl} = \lambda\delta_{ij}\epsilon_{kk} = \lambda \begin{pmatrix} \frac{\partial u_k}{\partial x_k} & 0 & 0 \\ 0 & \frac{\partial u_k}{\partial x_k} & 0 \\ 0 & 0 & \frac{\partial u_k}{\partial x_k} \end{pmatrix} = \begin{pmatrix} p & 0 & 0 \\ 0 & p & 0 \\ 0 & 0 & p \end{pmatrix} = p\delta_{ij} \quad (2.17)$$

Die Bewegungsgleichung (Gl. 2.13) vereinfacht sich zu

$$\rho \frac{\partial v_i}{\partial t} = \frac{\partial p}{\partial x_i} \quad (2.18)$$

die zusammen mit der zeitlichen Ableitung des Materialgesetzes (Gl. 2.17)

$$\frac{\partial p}{\partial t} = \lambda \frac{\partial \epsilon_{rr}}{\partial t} \quad (2.19)$$

die Spannungs-Geschwindigkeits Formulierung für die Behandlung mit finiten Differenzen ergibt.

2.4. Dämpfung

In der Seismik umfasst die Dämpfung alle Mechanismen, die zu einer Abnahme der Amplitude seismischer Wellen führt. Dazu zählen geometrische Dämpfung, Streuung und intrinsische Dämpfung. Lediglich bei letztgenanntem kommt es zu einer Energieumwandlung in Wärme. Beim geometrischen Ausbreitungsverlust und der Streuung an Inhomogenitäten verteilt sich die potentiell und kinematisch gespeicherte Energie auf ein größeres Volumen, wodurch die Energiedichte und damit die Amplituden abnehmen. Eine Wellenfront einer

Punktquelle breitet sich in drei Dimensionen auf einer Kugeloberfläche und in zwei Dimensionen auf einem Kreisumfang aus, weshalb die Energiedichte mit $\frac{1}{r^2}$ bzw. $\frac{1}{r}$ abnimmt. Die Energiedichte E ist proportional zur quadratischen Partikelauslenkung A^2 . Somit gilt $A_{3D} \sim \frac{1}{r}$ und $A_{2D} \sim \frac{1}{\sqrt{r}}$.

Die intrinsische Dämpfung lässt sich durch einen reellen Exponenten $\alpha(\omega)$ einer eindimensionalen Welle mit Kreisfrequenz ω und Ausbreitungsgeschwindigkeit $c(\omega)$ beschreiben: (Shearer, 2009)

$$A(x, t) = A_0 e^{-i\omega(t - \frac{x}{c(\omega)})} e^{-\alpha(\omega)x} \quad (2.20)$$

Eine anschauliche Bedeutung ergibt sich mit dem Qualitätsfaktor $Q(\omega)$, der zum Kehrwert des Energieanteils $\frac{\Delta E}{E}$ proportional ist, welcher bei jeder Schwingung in Wärme umgewandelt wird.

$$\frac{1}{Q(\omega)} = -\frac{\Delta E}{2\pi E} = \frac{2c\alpha(\omega)}{\omega} \quad (2.21)$$

Die Forderung nach einer kausalen Wellenausbreitung (Dehnung folgt zeitlich auf Druckänderung) impliziert eine frequenzabhängige Dämpfung $\alpha(\omega)$ sowie eine frequenzabhängige (dispersive) Ausbreitungsgeschwindigkeit $c(\omega)$ (Aki and Richards, 2002). Als erste Näherung kann Q über ein weites Frequenzband als konstant angesehen werden. Mit $\omega = 2\pi f$, $c = f\lambda$ und Gleichung 2.21 folgt:

$$e^{-\alpha(\omega)x} = e^{-\frac{\pi f x}{cQ}} = e^{-\frac{\pi x}{Q\lambda}} \quad (2.22)$$

Hieran ist zu erkennen, dass für $Q = 300$ die Amplitude nach $x \approx 100\lambda$ auf $\frac{1}{e}$ abgefallen und die Dämpfung α proportional zur Frequenz f ist.

2.5. Oberflächenwellen - Grenzflächenwellen

Grenzflächenwellen sind grundsätzlich von den Raumwellen zu unterscheiden. Sie breiten sich entlang von, durch unterschiedliche Materialparameter gebildete, Grenzflächen aus. Ein prominentes Beispiel ist durch die Erdoberfläche gegeben, weshalb oft von Oberflächenwellen gesprochen wird. Hieraus folgt bereits, dass die geometrische Dämpfung mit $\frac{1}{r}$ stattfindet und damit die Amplitudenabnahme proportional zu $\frac{1}{\sqrt{r}}$ ist (vgl. Kapitel 2.4).

Die Unterscheidung der Grenzflächenwellen erfolgt anhand ihrer Polarisierung. Love-Wellen

sind horizontal polarisierte Scherwellen (SH), Rayleigh-Wellen bestehen aus Longitudinalwellen und vertikal polarisierten Scherwellen (P-SV). Beide Typen entstehen durch an der Oberfläche interferierende Raumwellen. Die mathematische Herleitung erfolgt über spezielle Lösungen der Wellengleichung mit Randbedingungen. Es ergibt sich eine von der Frequenz abhängende Ausbreitungsgeschwindigkeit. Dieses $c(\omega)$ wird als geometrische Dispersionsbeziehung bezeichnet, da sie ihren Ursprung in den geometrischen Randbedingungen hat (vgl. Dispersion von Raumwellen in Kap. 2.4). Aus diesem Kapitel geht auch hervor, dass die Eindringtiefe seismischer Wellen von ihrer Frequenz abhängt. Bei Medien mit tiefenabhängigen Materialparametern folgt damit, dass sich die Frequenzen der Oberflächenwellen mit unterschiedlichen Geschwindigkeiten ausbreiten. Ein weiteres Phänomen ergibt sich aus den Eigenschaften des funktionalen Zusammenhangs $c(\omega)$. Mehrdeutigkeiten in den Lösungen führen zu mehreren Lösungsästen die als Moden bezeichnet werden und als einzelne Kurven in einem $c - \omega$ -Diagramm sichtbar werden (vgl. Gl. 2.23).

Geführte Wellen sind eine weitere Art von Wellen, die entlang von Grenzflächen gemessen werden können, jedoch nicht zu klassischen Oberflächenwellen gezählt werden, da sie sich nicht entlang einer Grenzfläche ausbreiten. Es handelt sich dabei um Raumwellen, die zwischen zwei Grenzflächen propagieren und durch Reflexion darin gefangen sind. Sie sind häufig in marinen seismischen Daten mit geringer Wassertiefe zu beobachten. Pekeris (1948) leitete erstmals die geometrische Dispersionsbeziehung von geführten Wellen in Flüssigkeitsschichten her. Die Erweiterung auf einen elastischen Meeresboden erfolgte durch Press and Ewing (1950).

$$\tan \left(K_n H \sqrt{\frac{c^2}{v_1^2} - 1} + n\pi \right) = \frac{\rho_2 \beta_2^4 \sqrt{\frac{c^2}{v_1^2} - 1}}{\rho_1 c^4 \sqrt{1 - \frac{c^2}{\alpha_2^2}}} \left[4 \sqrt{1 - \frac{c^2}{\alpha_2^2}} \sqrt{1 - \frac{c^2}{\beta_2^2}} - \left(2 - \frac{c^2}{\beta_2^2} \right)^2 \right] \quad (2.23)$$

H : Mächtigkeit der Flüssigkeitsschicht

ρ_1, ρ_2 : Dichte der Flüssigkeit und des Meeresbodens

v_1, α_2 : P-Wellengeschwindigkeit der Flüssigkeit und des Meeresbodens

β_2 : Scherwellengeschwindigkeit des Meeresbodens

Diese Gleichung definiert einen impliziten Zusammenhang zwischen Kreiswellenzahl K_n bzw. Kreisfrequenz ω_n ($\omega_n = cK_n$) und Phasengeschwindigkeit c . Aufgrund der π -Periodizität der Tangens-Funktion, ausgedrückt durch den Term $+n\pi$, können jedem K n Phasengeschwindigkeiten zugeordnet werden, wobei n die Mode der geführten Welle bezeichnet.

2.6. Slant-Stack

Ein slant-stack ist eine lineare Transformation einzelner Seismogrammspuren $u(x, t)$ von der x - t -Ebene in die p - τ -Ebene. Ziel ist eine Beschreibung der Intensität eines Wellenfeldes $u(x, t)$ in Abhängigkeit von Phasenlangsamkeit p ($p = \frac{1}{v}$ mit v : Phasengeschwindigkeit) und Interceptzeit τ .

Slant-stack bedeutet schräger Stapel, was eine intuitive Beschreibung der Methode darstellt. Die Amplitude jeder der N Spuren wird entlang der Transformationsgeraden $t = \tau + px$ aufsummiert, wodurch Wellenzüge mit gleichem p und τ konstruktiv interferieren. Ein slant-stack ist definiert als

$$\bar{u}(p, \tau) = \sum_{l=1}^N u(x_l, \tau + px_l) \quad (2.24)$$

Oberflächenwellen lassen sich aufgrund deren dispersiven Eigenschaften in einem p - ω (Langsamkeits-Frequenz)-Spektrum anhand der ausgebildeten Moden identifizieren. Das p - ω -Spektrum folgt aus einer Fourier-Transformation des slant-stacks vom τ - in den ω -Bereich. (McMechan and Yedlin, 1981)

$$\tilde{\bar{u}}(p, \omega) = \int_{-\infty}^{\infty} \sum_{l=1}^N u(x_l, \tau + px_l) e^{i\omega\tau} d\tau \quad (2.25)$$

Für die Anwendung ist es günstig eine mathematische Umformung vorzunehmen. Die Rücktransformation mit $\tau = t - px$ ergibt

$$\tilde{\bar{u}}(p, \omega) = \int_{-\infty}^{\infty} \sum_{l=1}^N u(x_l, t) e^{i\omega\tau} e^{-i\omega px_l} dt \quad \text{mit } \frac{d\tau}{dt} = 1 \quad (2.26)$$

Hier ist die Integration als Fourier-Transformation zu identifizieren, womit sich das p - ω -Spektrum aus der Summe der $\tilde{u}(x_l, \omega)$ berechnet, deren Phase jeweils um $e^{-i\omega px_l}$ verschoben wird.

$$\tilde{\bar{u}}(p, \omega) = \sum_{l=1}^N \tilde{u}(x_l, \omega) e^{-i\omega px_l} \quad (2.27)$$

3. Finite Differenzen Methode

Die Finite Differenzen Methode (FDM) ist ein numerisches Verfahren zur Lösung von gewöhnlicher und partieller Differentialgleichungen. Ausgangspunkt sind die drei Gleichungen 2.18, 2.17 und 2.5. Diese lassen sich nur für sehr einfache Fälle analytisch lösen. Komplizierte Randbedingungen und Modellvorgaben verlangen nach einem numerischen Ansatz.

Die Idee der FDM besteht darin, von der kontinuierlichen und analytischen Formulierung der PDGLs auf endlich bestimmte (finite) Differenzen überzugehen. Dazu wird die Infinitesimalrechnung gewissermaßen wieder rückgängig gemacht. Weiterhin wird Raum und Zeit diskretisiert, d.h. Punkte erstellt denen die relevanten Eigenschaften $(p, v_i, v_p, v_s, \rho, \dots)$ zugeschrieben werden. Die zeitliche Änderung einer Größe f in einem Punkt $\frac{\partial f}{\partial t}$ wird so durch die Änderung zwischen zwei Zeitschritten $\frac{f(x,t)-f(x,t+\Delta t)}{\Delta t}$ genähert. Analog wird der Gradient in einem Punkt $\frac{\partial f}{\partial h}$ durch die Änderung zwischen zwei Raumpunkten $\frac{f(x,t)-f(x+\Delta x,t)}{\Delta h}$ genähert. Statt der gesuchten Funktion $f(x, t)$ wird nun lediglich der Funktionswert an Stützstellen benötigt.

3.1. Diskretisierung von Raum und Zeit

Die für die Anwendung der FD-Operatoren benötigten Stützstellen, werden durch Aufteilung von Raum und Zeit in äquidistante Punkte generiert. Das Gitter hat eine Gesamtgröße von $NX \times NY$. Die Simulationsdauer beträgt $NT \cdot \Delta t$.

Zeit: $t = n \cdot \Delta t$	Zeitindex n , Zeitschritt Δt
Raum: $x = k \cdot \Delta x$	Raumindex k , Gitterpunktabstand Δx
$y = l \cdot \Delta y$	Raumindex l , Gitterpunktabstand Δy

Jedem Raumpunkt werden die relevanten zeitunabhängigen Materialparameter wie Dichte $\rho(x, y)$, sowie die zeitabhängigen Wellenfeldgrößen wie die Partikelgeschwindigkeit $v_x(x, y, t)$, welche die Wellenausbreitung beschreiben, zugeschrieben.

3.2. Definition des *Standard Staggered Grid* (SSG)

Das SSG (Virieux, 1986; Levander, 1988) ist eine Methode die physikalischen Größen so auf und zwischen Gitterpunkten zu verteilen, damit sich die partiellen Differentialgleichungen der Wellengleichung jeweils auf einen gemeinsamen Punkt im Gitter beziehen. Die Gestalt des SSG in zwei Dimensionen ist in Abbildung 3.1 dargestellt. Die zeitliche Dimension ist nicht explizit gezeigt. Der Druck p liegt zeitlich auf halben Gitterpunkten, die Partikelgeschwindigkeiten v auf ganzen.

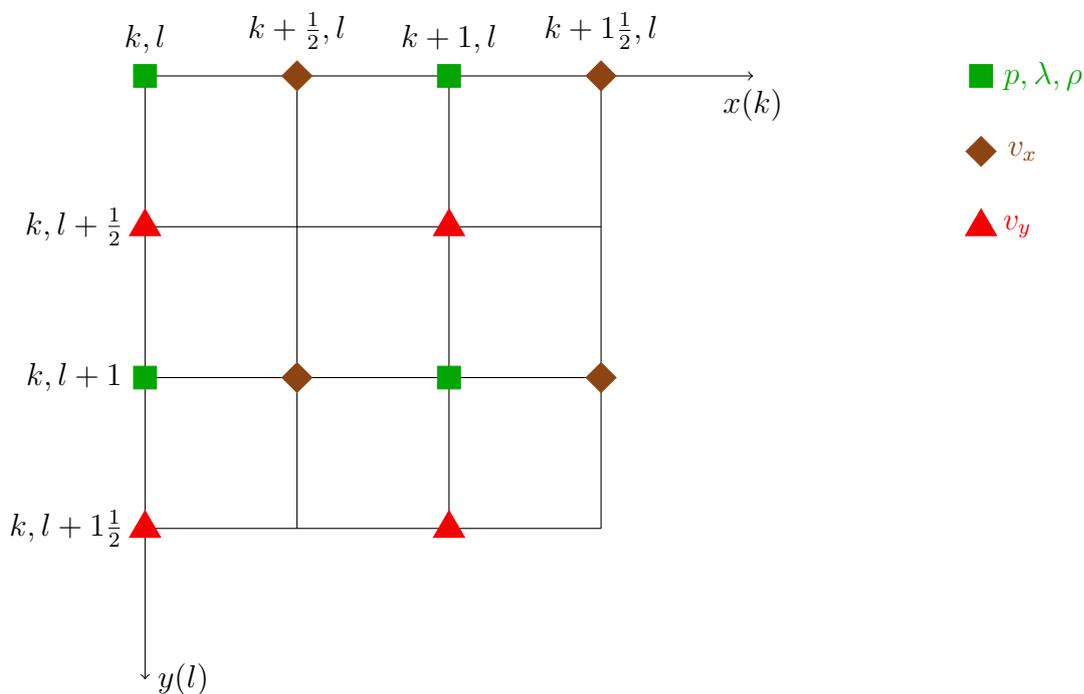


Abbildung 3.1.: Darstellung des *Standard Staggered Grid* (SSG) mit der Verteilung von Material- und Wellenfeldparametern in 2 Dimensionen.

Nach dem Lösungsansatz mit finiten Differenzen werden die analytischen Berechnungen der räumlichen und zeitlichen Ableitungen durch ein gewichtetes Mittel (FD-Operatoren) von Werten auf umgebenden Gitterpunkten genähert. Es stehen FD-Operatoren verschiedener Ordnung zur Verfügung, die eine unterschiedliche Anzahl von Gitterpunkten für die Näherung miteinbeziehen. Je weiter sie greifen, desto besser ist die Näherung, jedoch steigt auch der Berechnungsaufwand.

Für eine praktikable Umsetzung des FD-Schemas erfolgt die Lösung der akustischen Wel-

lengleichung in einem rekursiven Zeitschrittverfahren. Dazu werden die zeitlichen Ableitungen von Druck und Partikelgeschwindigkeiten durch FD-Operatoren 2. Ordnung ersetzt. Hierbei entstehen je zwei zeitlich aufeinander folgende Werte von p bzw. v_i .

$$\rho \frac{\partial v_i}{\partial t} = \frac{\partial p}{\partial x_i} \quad \text{mit} \quad \left. \frac{\partial v_i}{\partial t} \right|^{n+\frac{1}{2}} \approx \frac{v_i^{n+1} - v_i^n}{\Delta t}$$

$$\rightarrow v_i^{n+1} \approx v_i^n + \frac{\Delta t}{\rho} \left. \frac{\partial p}{\partial x_i} \right|^{n+\frac{1}{2}} \quad (3.1)$$

$$\frac{\partial p}{\partial t} = \lambda \frac{\partial \epsilon_{rr}}{\partial t} \quad \text{mit} \quad \left. \frac{\partial p}{\partial t} \right|^n \approx \frac{p^{n+\frac{1}{2}} - p^{n-\frac{1}{2}}}{\Delta t}$$

$$\rightarrow p^{n+\frac{1}{2}} \approx p^{n-\frac{1}{2}} + \Delta t \cdot \lambda \left. \frac{\partial \epsilon_{rr}}{\partial t} \right|^n \quad (3.2)$$

An Gleichung 3.1 ist bereits zu erkennen, wie die Definition der Geschwindigkeit auf ganzen und der Spannungen auf halben Gitterpunkten dazu führt, dass sich die Gleichungen auf einen gemeinsamen Punkt beziehen.

Eine kompakte Beschreibung des kompletten FD-Schemas lässt sich durch das Verwenden von Platzhaltern der räumlichen Ableitungen für FD-Operatoren beliebiger Ordnung erzielen. Hochgestellt findet sich der Zeitschritt und tiefgestellt der räumliche Mittelpunkt. Für 2. bzw. 4. Ordnung gilt:

$$\left. \frac{\partial p}{\partial x} \right|_{k+\frac{1}{2},l}^{n+\frac{1}{2}} \stackrel{2.O.}{=} \frac{p_{k+1,l}^{n+\frac{1}{2}} - p_{k,l}^{n+\frac{1}{2}}}{\Delta x} \quad (3.3)$$

$$\left. \frac{\partial p}{\partial x} \right|_{k+\frac{1}{2},l}^{n+\frac{1}{2}} \stackrel{4.O.}{=} \frac{\frac{9}{8} \left(p_{k+1,l}^{n+\frac{1}{2}} - p_{k,l}^{n+\frac{1}{2}} \right) - \frac{1}{24} \left(p_{k+2,l}^{n+\frac{1}{2}} - p_{k-1,l}^{n+\frac{1}{2}} \right)}{\Delta x} \quad (3.4)$$

Damit lässt sich die akustische Wellengleichung unter Verwendung des SSG wie folgt diskretisieren:

$$p_{k,l}^{n+\frac{1}{2}} = p_{k,l}^{n-\frac{1}{2}} + \Delta t \cdot \lambda_{k,l} \left(\left. \frac{\partial v_x}{\partial x} \right|_{k,l}^n + \left. \frac{\partial v_y}{\partial y} \right|_{k,l}^n \right)$$

$$v_{x,k+\frac{1}{2},l}^{n+1} = v_{x,k+\frac{1}{2},l}^n + \frac{\Delta t}{\rho_{k+\frac{1}{2},l}} \left. \frac{\partial p}{\partial x} \right|_{k+\frac{1}{2},l}^{n+\frac{1}{2}}$$

$$v_{y,k,l+\frac{1}{2}}^{n+1} = v_{y,k,l+\frac{1}{2}}^n + \frac{\Delta t}{\rho_{k,l+\frac{1}{2}}} \left. \frac{\partial p}{\partial y} \right|_{k,l+\frac{1}{2}}^{n+\frac{1}{2}}$$

Die Verwendung des SSG erfordert hinsichtlich numerischer Stabilität, dass die Dichte ρ auf Zwischengitterpunkten aus den Dichtewerten der umgebenden ganzzahligen Gitterpunkten arithmetisch gemittelt wird.

$$\rho_{k+\frac{1}{2},l} = \frac{\rho_{k,l} + \rho_{k+1,l}}{2} \quad \rho_{k,l+\frac{1}{2}} = \frac{\rho_{k,l} + \rho_{k,l+1}}{2}$$

3.3. Ablauf der Wellenfeldberechnung

Zur Simulation des Wellenfeldes in Raum und Zeit muss jede der 3 diskretisierten Gleichungen an jedem Ort und zu jeder Zeit ausgewertet werden. Da jeweils der zeitlich vorangegangene Wert benötigt wird muss n stetig erhöht werden. Dagegen spielt die Reihenfolge der räumlichen Abarbeitung keine Rolle. Dagegen wird der Druck $p^{n+\frac{1}{2}}$ für die Berechnung der Partikelgeschwindigkeiten $v_i^{n+\frac{1}{2}}$ benötigt und müssen daher zuerst berechnet werden.

```

for  $n=1$  to  $NT$  do
  for  $k=1$  to  $NX$  do
    for  $l=1$  to  $NY$  do
       $\sqsubset$  Berechne Druck  $p$ 
    for  $k=1$  to  $NX$  do
      for  $l=1$  to  $NY$  do
         $\sqsubset$  Berechne Partikelgeschwindigkeiten  $v_i$ 

```

Algorithm 1: Finite Differenzen Simulationsablauf

3.4. Wahl der Diskretisierung

Die Wahl der räumlichen (Δh) und zeitlichen (Δt) Schrittweite bestimmt neben der Ordnung der FD-Operatoren maßgeblich die Güte und Stabilität der Simulation. Bei zu großen Werten kommt es zu ausgeprägter Gitterdispersion, sowie numerischer Instabilität, wodurch die Ergebnisse unbrauchbar werden. Zu kleine Werte führen zu einem enormen Rechenaufwand. Daher existieren Regeln für die Wahl von Δh und Δt .

3.4.1. Räumliche Abtastung - Gitterdispersion

Der Gitterabstand Δnh kann nur dann sinnvoll bewertet werden, wenn er in Bezug zu der räumlichen Länge der vorhandenen Signale gesetzt wird. Diese ist durch die Wellenlänge λ gegeben. Die Frage ist nun einerseits wie viele Stützstellen innerhalb einer Wellenlänge liegen müssen, damit die Differenzenquotienten dieser Stützstellen zu einer hinreichend genauen Näherung der analytisch exakten Ableitung führen. Andererseits ist zu klären was hinreichend bedeutet bzw. was im Falle unzureichender Abtastung passiert.

Als sehr anschauliches Beispiel sei hier auf [Bohlen et al. \(2015\)](#) verwiesen, wo als repräsentative Wellenform $v_x = \sin\left(2\pi\frac{x}{\lambda}\right)$ herangezogen wird. Deren analytisch und numerisch, mit dem FD-Operator 2-ter Ordnung für verschiedene Gitterabstände Δh , berechnete Ableitung an der Stelle $x = 0$ ist in Tabelle 3.1 dargestellt. Der FD-Operator 2-ter Ordnung lautet:

$$\left.\frac{dv_x}{dx}\right|_{x=0} \approx \left.\frac{v_x(x + \Delta x) - v_x(x)}{\Delta x}\right|_{x=0} \quad (3.5)$$

Die Ergebnisse in Tabelle 3.1 geben einen ersten Eindruck von der Genauigkeit der Ableitung, doch sind deren Wirkung in einer FD-Simulation damit nicht abzusehen. Zu diesem Zweck wird hier auf Abbildung 3.2 hingewiesen. Sie zeigt die Ergebnisse einer FD-Simulation mit variierter Anzahl n Stützstellen pro minimaler Wellenlänge. Bei $n = 16$ ist das Ende der Welle stets scharf abgegrenzt. Bei $n = 4$ und besonders bei $n = 2$ verschmiert das Wellenende mit der Ausbreitung zunehmend in einer Form, wie sie physikalisch durch Dispersion hervorgerufen wird. Es handelt sich im Speziellen um normale Dispersion, da die Ausbreitungsgeschwindigkeit mit zunehmender Wellenlänge wächst. Hier hat die Dispersion jedoch keine physikalische Ursache, sondern ist vielmehr ein numerisches Artefakt aufgrund unzureichender räumlicher Diskretisierung und wird daher Gitterdispersion genannt.

Tabelle 3.1.: Vergleich der analytisch und numerisch (Gl.: 3.5), mit verschiedenen Gitterabständen $\Delta h = \frac{\lambda}{n}$, berechneten Ableitung von $v_x = \sin\left(2\pi\frac{x}{\lambda}\right)$.

n	Δh	$\frac{dv_x}{dx} _{x=0}$
analytisch	-	$2\pi \approx 6.283$
2	$\lambda/2$	0.0
4	$\lambda/4$	4.0
8	$\lambda/8$	5.657
16	$\lambda/16$	6.123

Durch das Einhalten des folgenden Kriteriums wird die Gitterdispersion auf einem für die meisten Anwendungen akzeptablem Maß gehalten. Bei n ist die Abhängigkeit von der Länge des FD-Operators zu beachten (s. dazu Tabelle 3.2).

$$\Delta h \leq \frac{\lambda_{\min}}{n} = \frac{v_{s,\min}}{nf_{\max}} \quad (3.6)$$

$v_{s,\min}$: kleinste im Modell vorkommende v_s -Geschwindigkeit

f_{\max} : höchste Frequenz im Quellsignal

n : Anzahl der Stützstellen pro minimaler Wellenlänge

Tabelle 3.2.: Anzahl n der benötigten Gitterpunkte pro minimaler Wellenlänge für verschiedene Operatorlängen und -typen, um die Gitterdispersion weitgehend zu unterdrücken.

FD-Ordnung	n (Taylor)	n (Holberg)
2	12	12
4	8	8.32
6	7	4.77
8	6	3.69
10	5	3.19
12	4	2.91

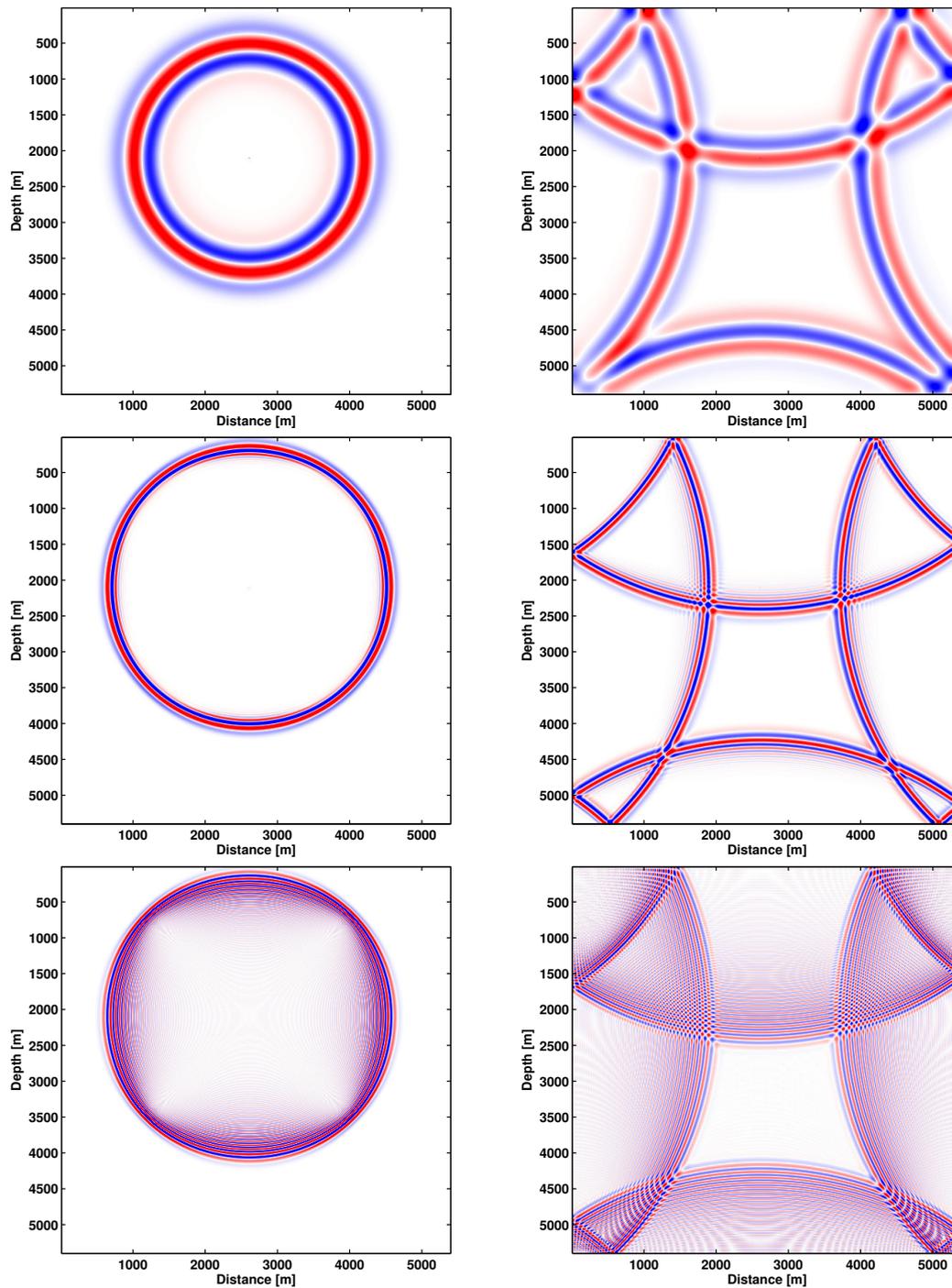


Abbildung 3.2.: Gestalt der Gitterdispersion mit zunehmender Propagation (links -> rechts) bei Verwendung von 16 (oben), 4 (Mitte) und 2 (unten) Gitterpunkten pro minimaler Wellenlänge aus (Bohlen et al., 2015).

3.4.2. Zeitliche Abtastung - Numerische Instabilität

Um numerische Instabilität zu vermeiden muss Δt kleiner sein als die Ausbreitungsdauer einer Welle zwischen zwei benachbarten Gitterpunkten $\left(\frac{\Delta h}{v_{p,\max}}\right)$. Das Courant-Friedrichs-Lewy-Kriterium besagt hierzu:

$$\Delta t \leq \frac{\Delta h}{\gamma v_{p,\max} \sqrt{d}}$$

$v_{p,\max}$: größte im Modell vorkommende v_p -Geschwindigkeit

γ : Konstante für FD-Ordnung s. Tabelle 3.3

d : räumliche Modelldimension 2-D $\rightarrow d = 2$

Tabelle 3.3.: Konstante γ des Courant-Friedrichs-Lewy-Kriteriums für die FD-Ordnungen 2 bis 12.

FD-Ordnung	γ (Taylor)	γ (Holberg)
2	1	1
4	$\frac{7}{6}$	1.184614
6	$\frac{149}{120}$	1.283482
8	$\frac{2161}{1680}$	1.345927
10	$\frac{53089}{40320}$	1.38766
12	$\frac{1187803}{887040}$	1.417065

Die zeitliche Diskretisierung führt stets auch zu numerischer Dispersion. Wird das Courant-Friedrichs-Lewy-Kriterium verletzt, kommt es zur Instabilität, d.h. die Amplituden wachsen uneingeschränkt.

4. Wellenforminversion

Dieses Kapitel soll als verständliche und vollständige Beschreibung der theoretischen Grundlagen der Wellenforminversion dienen. Nach einer Erläuterung der Problemstellung aus unterschiedlichen Perspektiven wird auf die Herausforderungen der nichtlinearen Inversion eingegangen. Die Thematik der Abstiegsrichtung und der Schrittweite werden als zentrale Aspekte ausführlich behandelt.

4.1. Geophysikalische Perspektive

Unter der Wellenforminversion (*FWI: full waveform inversion*) versteht man die Anstrengung aus punktuell und zeitlich aufgezeichneten Wellenzügen, z.B. mit Seismometern, die räumliche Verteilung der Materialeigenschaften, durch die sich die Wellen ausgebreitet haben, zu rekonstruieren. Es handelt sich um eine Umkehrung (Inversion) des Vorwärtsproblems

$$\vec{u} = G(\vec{m}) \tag{4.1}$$

mit Wellenfeld \vec{u} , Vorwärtsoperator G (Wellengleichung) und Modell \vec{m} . Das Vorwärtsproblem lässt sich durch numerisches Lösen der Wellengleichung, beispielsweise mit finiten Differenzen, für beliebige Modelle behandeln (s. Kapitel 3). Dazu ist es notwendig das Modell zu diskretisieren, d.h. es räumlich in endliche Blöcke aufzuteilen, welche die Modellparameter zugeordnet bekommen. Je nach Problemstellung bestehen die Modellparameter in jedem Block aus bis zu 22 (allgemeiner Fall: linear elastisch, anisotrop) oder nur aus 2 Materialparameter (akustisch, isotrop). Das Wellenfeld \vec{u} enthält für jeden Modellblock den zeitlichen Verlauf des Drucks und der Partikelgeschwindigkeiten (zeitlich diskretisiert).

Schwieriger gestaltet sich die Umkehrung der Fragestellung (Inversion), d.h. die Suche nach einem Modell \vec{m} , welches das Wellenfeld \vec{u} , unter gegebenen Anfangsbedingungen (Quellsignal zum Zeitpunkt t_0) hervorruft.

4.2. Mathematische Perspektive

Eine nähere Beschreibung auf abstrakter Ebene hilft die Problemstellung besser einordnen zu können. Die Wellenforminversion kann als *kombinatorisches Optimierungsproblem* angesehen werden. Die Aufgabe besteht darin, aus der Menge aller möglichen Kombinationen L der Modellparameter, eine Variante ψ zu finden, sodass eine zugeordnete Zielfunktion $S(\psi)$ einen möglichst kleinen Wert annimmt. Die Größe des Lösungsraumes L ist unter der Annahme eines diskretisierten 2D-Modells mit $NX \times NY$ (100×100) Elementen und geologisch realistischen Parameterwerten zwischen m_{\min} ($1000 \frac{\text{m}}{\text{s}}$) und m_{\max} ($5000 \frac{\text{m}}{\text{s}}$) in quantisierten Schritten ($10 \frac{\text{m}}{\text{s}}$) endlich (hier für einen Materialparameter). Damit berechnet sich die Größe des Lösungsraumes als die Variation mit Wiederholung von 10 000 Elementen aus 400 Objekten unter Beachtung der Reihenfolge mit 400^{10000} .

4.2.1. Zielfunktion

Zur Bewertung jeder Lösung ψ dient die Zielfunktion $S(\psi)$. Sie quantifiziert den Abstand zwischen den observierten Daten \vec{d} und den synthetischen Daten \vec{u} . Zur Erzeugung der synthetischen Daten ist für jede Lösung ψ eine Wellenfeldsimulation notwendig. Als physikalisch sinnvolles Fehlermaß dient die L2-Norm (L2-Misfit)

$$S = \frac{1}{2} \sum_{i=1}^N (d_i - u_i)^2 \quad (4.2)$$

Die Datenresiduen $d_i - u_i$ gehen quadratisch ein und da die Messgröße Partikelgeschwindigkeit oder Druck darstellt, sind die Datenresiduen proportional zur Residuumsenergie. Anschaulich ist dies die Energie, die durch die betrachtete Lösung nicht erklärt werden kann. Der Faktor $\frac{1}{2}$ ist Konvention und wird eingeführt, damit er bei späteren Ableitungen den Faktor 2 weghebt.

4.2.2. Lösungsstrategie

Für viele Problemstellungen aus dem Bereich der kombinatorischen Optimierung ist es erforderlich eine individuelle Lösungsstrategie zu entwickeln, bzw. bestehende Konzepte

anzupassen. Maßgebendes Kriterium ist meist die erforderliche Rechenzeit. So scheidet beispielsweise die universelle *Trial and Error* oder *Brute-Force* Methode aufgrund des dramatisch großen Lösungsraums aus.

Für die FWI hat sich die *lokale Suche* als praktikables metaheuristisches Suchverfahren etabliert. Hierbei wird ausgehend von einem Startmodell iterativ eine gemäß der Zielfunktion bessere Lösung in der Umgebung der aktuellen gesucht.

Die Hadamard'schen Forderungen zeigen die fundamentalen Schwierigkeiten denen man aus mathematischer Sicht gegenübersteht. Sind sie erfüllt spricht man von einem korrekt gestellten Problem (Hadamard, 1902).

1. Existenz - Das Problem besitzt eine Lösung.
2. Eindeutigkeit - Die Lösung ist eindeutig.
3. Stabilität - Die Lösung hängt stetig von den Eingangsdaten ab.

Man sollte sich zunächst über die Bedeutung der Forderungen klar werden. Dass eine Lösung existiert scheint klar zu sein, doch ist dies nicht streng gültig, wenn man bedenkt, dass die aufgezeichneten Seismogramme Messfehler wie Rauschen enthalten. Eine gefundene Lösung ist im Allgemeinen nicht eindeutig. Dies kann schnell eingesehen werden. So können Modellblöcke derart am Rand des Messgebiets liegen, dass die aufgezeichneten Wellen nicht durch diese hindurchgelaufen sind, womit die Lösung für diese Bereiche keinen Einfluss hat.

4.3. Lineare und nicht lineare Inversionsprobleme

Unter nicht linearen Problemen wird ein Zusammenhang zwischen Messdaten d_i und Modellparametern m_i verstanden, welcher selbst von den Modellparametern abhängt. Zudem gehen die m_i nicht linear ein. Ganz allgemein gilt $\vec{u} = G(\vec{m})$. Beispiele für solche Probleme sind Geoelektrik, elektromagnetische Methoden und die FWI.

Bei einem linearen Problem kann eine vereinfachte Matrix-Vektormultiplikation verwen-

det werden:

$$\vec{u} = R \vec{m} \quad (4.3)$$

Daten und Modell sind über konstante und durchaus bekannte analytische Funktionen verknüpft. Alle m_i gehen linear ein. Beispiele hierfür sind Gravimetrie und Magnetik.

4.3.1. Inversion linearer Probleme

Bevor den nicht linearen Problemen und der FWI volle Aufmerksamkeit geschenkt wird, dient eine kurze Behandlung der linearen Probleme einem tieferen Verständnis. Es gilt $\vec{u} = R \vec{m}$ mit $R \in \mathbb{R}^{N \times M}$, $\vec{u} \in \mathbb{R}^N$ und $\vec{m} \in \mathbb{R}^M$. N gibt die Anzahl der Gleichungen (Messungen) an. Sie ergibt sich aus dem Produkt aus der Anzahl von Schüssen und Empfängern. M ist die Anzahl der Modellblöcke, was bei der Inversion der Zahl der Unbekannten entspricht. Zur Lösung von $\vec{m} = R^{-1} \vec{u}$ wäre es aus mathematischer Sicht wünschenswert, dass $N = M$ gilt und R regulär ist, d.h. vollen Rang besitzt (alle Zeilenvektoren linear unabhängig), dann wäre R leicht zu invertieren. Dazu müsste jeder Modellblock in Summe aller Messungen genau einmal beleuchtet worden sein. Da dies in der Praxis nie zutrifft, muss ein Weg gefunden werden um eine Inversion mit $N \neq M$ durchführen zu können. Dazu wird die Methode der kleinsten Fehlerquadrate vorgestellt.

4.3.2. Methode der kleinsten Fehlerquadrate

Ziel ist es $\vec{u} = R \vec{m}$ mit $R \in \mathbb{R}^{N \times M}$ und $N \neq M$ zu invertieren, d.h. nach \vec{m} aufzulösen, sodass die Summe der Abweichungen im Quadrat zwischen gemessenen Daten \vec{d} und den synthetischen Daten \vec{u} , wie sie sich mit dem als korrekt angesehenen Modell \vec{m} ergeben, minimal ist. Diese Abweichung zwischen \vec{d} und \vec{u} wird als Misfit S bezeichnet.

$$S = \sum_{i=1}^N \left(d_i - \sum_{j=1}^M R_{ij} m_j \right)^2 \quad (4.4)$$

$$S = (\vec{d} - R \vec{m})^T (\vec{d} - R \vec{m}) \quad (4.5)$$

Das Extremum (Minimum) von S bezüglich der Modellparameter m_q findet man durch Nullsetzen der Ableitung (Menke, 1989).

$$\frac{\partial S}{\partial m_q} = 2 \sum_{k=1}^m m_k \sum_{i=1}^n R_{iq} R_{ik} - 2 \sum_{i=1}^n R_{iq} d_i = 0 \quad (4.6)$$

$$\rightarrow \sum_{k=1}^m \sum_{i=1}^n R_{iq} R_{ik} m_k = \sum_{i=1}^n R_{iq} d_i \quad (4.7)$$

mit $R_{iq} R_{ik} m_k = R_{iq} d_i$ und $R_{qi}^T R_{ik} m_k = R_{qi}^T d_i$ folgt die Normalengleichung

$$R^T R \vec{m} = R^T \vec{d} \quad \rightarrow \quad \vec{m} = (R^T R)^{-1} R^T \vec{d} \quad (4.8)$$

Liegt ein lineares Inversionsproblem vor, so kann das optimale Modell \vec{m} im Sinne der kleinsten Fehlerquadrate mit Gleichung 4.8 berechnet werden.

4.4. Inversion nicht linearer Probleme

Ziel ist es auch hier die Datenabweichung $\Delta \vec{d}$ zwischen gemessenen und mit dem als wahr angesehenen Modell \vec{m} berechneten Daten \vec{u} , also $\Delta \vec{d} = \vec{d} - G(\vec{m})$, zu minimieren. Für den L2-Misfit bezüglich des Modells \vec{m} gilt $S(\vec{m}) = \frac{1}{2} \Delta \vec{d}^T \Delta \vec{d}$. Die Nichtlinearität in $G(\vec{m})$ bezüglich \vec{m} schlägt sich bis zum Misfit $S(\vec{m})$ durch. Um mit $S(\vec{m})$ weiter verfahren zu können, wird eine Taylorentwicklung bis 2. Ordnung durchgeführt. Damit ist der Misfit $\tilde{S}(\vec{m})$ eine Näherung. Wird nun ein Minimum des Misfits gesucht, kann dieses nicht mehr als analytischen Ausdruck angegeben werden. Vielmehr sollte eine neue Lösungsstrategie angewendet werden.

Da eine Taylorentwicklung nur in der Umgebung des Entwicklungspunktes hinreichend exakte Ergebnisse liefert, sollte das Minimum auch nur in dieser Umgebung gesucht werden. Um dies zu veranschaulichen kann das Minimierungsproblem abstrahiert werden indem angenommen wird, dass der Modellvektor aus nur einem Wert besteht (s. Abbildung 4.1).

Die Suche nach dem Minimum von $\tilde{S}(\vec{m})$ ist dank der Näherung nun durchführbar. Sie führt allerdings nicht direkt auf das Minimum von $S(\vec{m})$, sondern auf das lokale Minimum $\vec{m}_0 + \Delta \vec{m}$ der Näherung $\tilde{S}(\vec{m})$. $\Delta \vec{m}$ kann somit als Modellverbesserung angesehen werden,

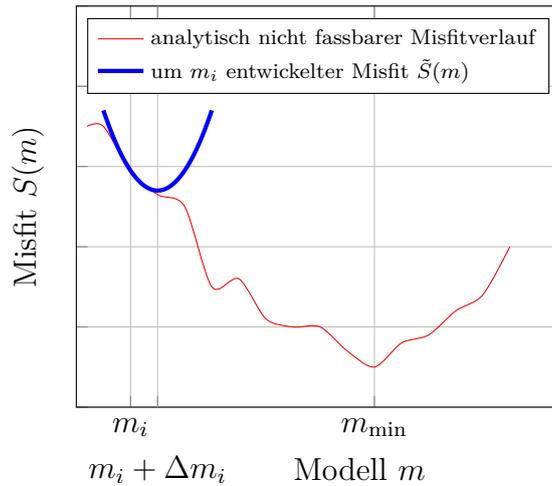


Abbildung 4.1.: Visualisierung des nicht linearen Misfitverlaufs, sowie die Entwicklung bis zur 2. Ordnung im Punkt m_i .

die das Modell \vec{m}_{\min} näher bringt. Es liegt nun nahe die lokale Minimierung iterativ anzuwenden um sich \vec{m}_{\min} allmählich zu nähern. Das Iterationsschema hat damit folgende Gestalt:

$$\vec{m}_{i+1} = \vec{m}_i + \Delta \vec{m}_i \quad (4.9)$$

Die Entwicklung der Misfitfunktion $S(\vec{m})$ hat um \vec{m}_i zu erfolgen. Die Taylorentwicklung einer skalarwertigen Funktion mit vektorwertigem Argument bis zur 2. Ordnung mit Entwicklungspunkt \vec{a} wird Schmiegequadratik genannt und kann kompakt geschrieben werden. $H_{f,\vec{x}}(\vec{a})$ ist die Hesse-Matrix von f bezüglich \vec{x} an der Stelle \vec{a} .

$$T_2 f(\vec{x}; \vec{a}) = f(\vec{a}) + \nabla f(\vec{a})^T (\vec{x} - \vec{a}) + \frac{1}{2} (\vec{x} - \vec{a})^T H_{f,\vec{x}}(\vec{a}) (\vec{x} - \vec{a}) \quad (4.10)$$

Damit gilt für die Entwicklung der Misfitfunktion $S(\vec{m})$ an der Stelle \vec{m}_i .

$$S(\vec{m}) = S(\vec{m}_i + \Delta \vec{m}_i) \approx S(\vec{m}_i) + \nabla_{\vec{m}} S(\vec{m}_i)^T \Delta \vec{m}_i + \frac{1}{2} \Delta \vec{m}_i^T H_{S,\vec{m}}(\vec{m}_i) \Delta \vec{m}_i + \mathcal{O}(\Delta \vec{m}_i^3) \quad (4.11)$$

Das Minimum wird über die Nullstelle der Ableitung von $S(\vec{m})$ nach der Modellverbesserung $\Delta \vec{m}_i$ bestimmt. Hierbei wird die Indexschreibweise verwendet. Indizes außerhalb von Klammern bezeichnen die Vektorkomponente und sind damit von dem Modell \vec{m}_i der i -ten Iteration zu unterscheiden. Es wird die Symmetrie von $H_{S,\vec{m}}(\vec{m}_i)$ verwendet, sowie

ausgenutzt, dass doppelte Indizes umbenannt werden dürfen.

$$\begin{aligned}
\left(\frac{\partial S(\vec{m})}{\partial \Delta \vec{m}_i}\right)_k &= (\partial \Delta \vec{m}_i)_k S(\vec{m}) \hat{=} \partial_k S(\vec{m}) \\
&\stackrel{4.11}{=} \partial_k \left(S(\vec{m}_i) + (\nabla_{\vec{m}} S(\vec{m}_i))_r (\Delta \vec{m}_i)_r + \frac{1}{2} (\Delta \vec{m}_i)_w (H_{S,\vec{m}}(\vec{m}_i))_{wl} (\Delta \vec{m}_i)_l \right) \\
&= (\nabla_{\vec{m}} S(\vec{m}_i))_r \delta_{kr} + \frac{1}{2} (H_{S,\vec{m}}(\vec{m}_i))_{wl} \partial_k ((\Delta \vec{m}_i)_w (\Delta \vec{m}_i)_l) \\
&= (\nabla_{\vec{m}} S(\vec{m}_i))_k + \frac{1}{2} (H_{S,\vec{m}}(\vec{m}_i))_{wl} (\delta_{kw} (\Delta \vec{m}_i)_l + (\Delta \vec{m}_i)_w \delta_{kl}) \\
&= (\nabla_{\vec{m}} S(\vec{m}_i))_k + \frac{1}{2} [(H_{S,\vec{m}}(\vec{m}_i))_{kl} (\Delta \vec{m}_i)_l + (H_{S,\vec{m}}(\vec{m}_i))_{wk} (\Delta \vec{m}_i)_w] \\
&\stackrel{H=H^T}{=} (\nabla_{\vec{m}} S(\vec{m}_i))_k + \frac{1}{2} [(H_{S,\vec{m}}(\vec{m}_i))_{kl} (\Delta \vec{m}_i)_l + (H_{S,\vec{m}}(\vec{m}_i))_{kw} (\Delta \vec{m}_i)_w] \\
&\stackrel{w \rightarrow l}{=} (\nabla_{\vec{m}} S(\vec{m}_i))_k + \frac{1}{2} [(H_{S,\vec{m}}(\vec{m}_i))_{kl} (\Delta \vec{m}_i)_l + (H_{S,\vec{m}}(\vec{m}_i))_{kl} (\Delta \vec{m}_i)_l] \\
&= (\nabla_{\vec{m}} S(\vec{m}_i))_k + (H_{S,\vec{m}}(\vec{m}_i))_{kl} (\Delta \vec{m}_i)_l = 0
\end{aligned}$$

In Vektorschreibweise ergibt sich

$$\frac{\partial S(\vec{m})}{\partial \Delta \vec{m}_i} = \nabla_{\vec{m}} S(\vec{m}_i) + H_{S,\vec{m}}(\vec{m}_i) \Delta \vec{m}_i = 0 \quad (4.12)$$

womit der Ausdruck für $\Delta \vec{m}_i$, bei dem das Minimum der quadratischen Näherung der i -ten Iteration erreicht wird, gefunden wurde.

$$\Delta \vec{m}_i = -H_{S,\vec{m}}(\vec{m}_i)^{-1} \nabla_{\vec{m}} S(\vec{m}_i) \quad (4.13)$$

Das ausgeschriebene Iterationsschema lautet

$$\vec{m}_{i+1} = \vec{m}_i - H_{S,\vec{m}}(\vec{m}_i)^{-1} \nabla_{\vec{m}} S(\vec{m}_i) \quad (4.14)$$

Betrachtet man das Iterationsschema, so sieht man, dass sich die Modellverbesserung \vec{m}_{i+1} aus dem vorangegangenen Modell \vec{m}_i ergibt, das entlang einer sogenannten Abstiegsrichtung modifiziert wird. Die sich hier ergebende spezielle Abstiegsrichtung $-H_{S,\vec{m}}(\vec{m}_i)^{-1} \nabla_{\vec{m}} S(\vec{m}_i)$ heißt Newton Richtung. Sie enthält aufgrund der Taylorentwicklung bis 2. Ordnung Terme mit Ableitungen 1. Ordnung (Gradient) und auch 2. Ordnung (Hesse-Matrix). Wie sich in nachfolgendem Kapitel zeigen wird, existieren diverse Möglichkeiten für die Wahl der Abstiegsrichtung, wobei das Gerüst der iterativen Modellverbesserung erhalten bleibt.

4.5. Abstiegsrichtung

Eine formale Definition der Abstiegsrichtung lautet (Gerds, 2007): $\vec{p} \in \mathbb{R}^n$ heißt Abstiegsrichtung von $S : \mathbb{R}^n \rightarrow \mathbb{R}$ in $\vec{m} \in \mathbb{R}^n$, falls es ein $\tilde{\alpha} > 0$ gibt mit

$$S(\vec{m} + \alpha\vec{p}) < S(\vec{m}) \quad \forall 0 < \alpha < \tilde{\alpha} \quad (4.15)$$

Anschaulich bedeutet dies, dass es eine maximale Schrittweite geben muss, für die sich der Misfit S , bei einer Modelländerung in Richtung \vec{p} , verkleinert. Eine hinreichende Bedingung ist eine negative Richtungsableitung von S in Richtung \vec{p} .

$$\nabla_{\vec{p}} S(\vec{m}) = \lim_{h \rightarrow 0} \frac{S(\vec{m} + h\vec{p}) - S(\vec{m})}{h} = \nabla S(\vec{m})^T \vec{p} < 0 \quad (4.16)$$

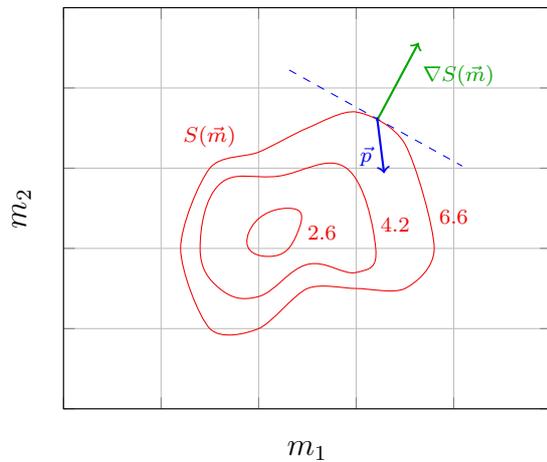


Abbildung 4.2.: Veranschaulichung der Definition 4.16 für die Abstiegsrichtung \vec{p} . Der Misfit $S(\vec{m})$ ist als Projektion über Höhenlinien abhängig von einem zweidimensionalen Modell dargestellt.

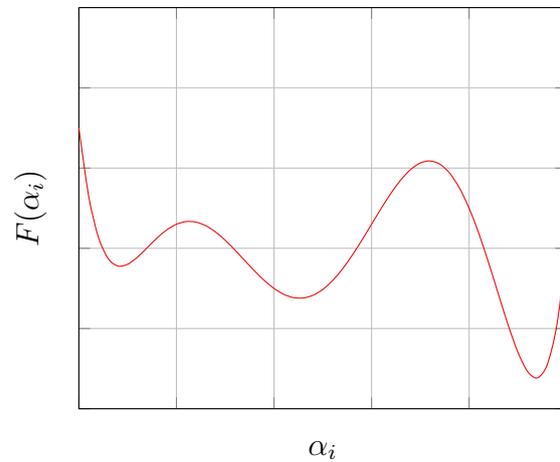


Abbildung 4.3.: Prinzipieller Verlauf des Misfits $F(\alpha_i)$, dargestellt in Abhängigkeit der Schrittweite α_i für eine fest gewählte Abstiegsrichtung \vec{p}_i .

Diese Bedingung wird bei Betrachtung von Abbildung 4.2 und Interpretation von $\nabla S(\vec{m})^T \vec{p}$ als Skalarprodukt verständlich. Der Gradient $\nabla S(\vec{m})$ steht in jedem Punkt \vec{m} senkrecht auf den Höhenlinien. Er zeigt in Richtung ansteigendem S . Beträgt der Winkel zwischen \vec{p} und $\nabla S(\vec{m})$ 90° bis 270° . So zeigt \vec{p} in Richtung Abstieg und Gleichung 4.16 ist erfüllt.

4.5.1. Berechnung des Gradienten

Der Gradient $\nabla_{\vec{m}} S(\vec{m}_i)$ der zu minimierenden Zielfunktion S bezüglich der Modellparameter \vec{m}_i ist grundlegender Bestandteil aller Abstiegsverfahren. Er muss in jeder Iteration i aktualisiert werden und ist damit neben der Bestimmung der Schrittweite eine zeitkritische Operation.

Der Gradient der Zielfunktion wird auch *sensitivity kernel* genannt, was seine anschauliche Bedeutung unterstreicht. Seine Komponenten $\nabla_{\vec{m}} S(\vec{m}_i) = (\partial m_1 S, \partial m_2 S, \dots, \partial m_n S)^T$ geben an, wie empfindlich die Zielfunktion auf eine Änderung der Modellparameter reagiert.

Eine offensichtliche Methode ist das Vorwärtsproblem G für das Modell \vec{m}_i zu lösen. Es resultieren die synthetischen Daten $\vec{u}_i = G(\vec{m}_i)$, die Datenabweichungen $\Delta \vec{d}_i = \vec{d}_i - \vec{u}_i$ und damit der Misfit $S(\vec{m}_i) = \frac{1}{2} \Delta \vec{d}_i^T \Delta \vec{d}_i$. Anschließend wird der Modellvektor in jeder Dimension einzeln um δm variiert. Damit ergeben sich n Modelle $\vec{m}_i^{\delta, j} = \vec{m}_i + \hat{e}_j \delta m$ mit $j = 1, 2, \dots, n$ für die alle eine Vorwärtsrechnung durchgeführt und der Misfit $S(\vec{m}_i^{\delta, j})$ berechnet wird. Nun kann der Gradient in jeder Dimension als Differenzenquotient (Sekantensteigung) genähert werden.

$$\nabla_{\vec{m}} S(\vec{m}_i) = \frac{1}{\delta m} \begin{pmatrix} S(\vec{m}_i) - S(\vec{m}_i^{\delta, 1}) \\ S(\vec{m}_i) - S(\vec{m}_i^{\delta, 2}) \\ \vdots \\ S(\vec{m}_i) - S(\vec{m}_i^{\delta, n}) \end{pmatrix} \quad (4.17)$$

In jeder Iteration i ist damit ein enormer Aufwand von $(n + 1)$ Vorwärtsrechnungen notwendig, weshalb ein effizienteres Verfahren benötigt wird.

4.5.2. Die adjoint state Methode

Die adjoint state Methode ermöglicht die effiziente Berechnung des Gradienten einer Zielfunktion im Zusammenhang eines numerischen Optimierungsproblems mit partiellen Differentialgleichungen (Wellengleichung) als Nebenbedingung, insbesondere für hochdimen-

sionale Modellräume. Der Aufwand ist durch das zweimalige Lösen der Wellengleichung mit speziellen Randbedingungen je Iteration gegeben. Dieses Teilproblem kann mit den gleichen Werkzeugen bearbeitet werden, die auch zur Berechnung der synthetischen Daten \vec{u}_i verwendet werden (z.B. FEM oder FDM).

Die Berechnung des Gradienten erfolgt entweder mit der Born Approximation und der Reziprozität der Green'schen Funktion (Tarantola, 1988) oder mit der adjoint state Methode bei der Lagrange-Multiplikatoren verwendet werden können (Plessix, 2006; Kamath and Tsvankin, 2016; Liu and Tromp, 2006; Bradley, 2013). Der Zugang gestaltet sich ohne spezifisches Vorwissen oftmals schwierig.

4.5.3. Konjugiertes Gradientenverfahren

Grundsätzlich lässt sich der reine Gradient als Abstiegsrichtung verwenden. Die Frage, die sich stellt, ist wie das Abstiegs- oder Konvergenzverhalten aussieht. Hierzu wird in diesem Kapitel gezeigt, dass zweckmäßige Modifikationen des Gradienten möglich sind, die zu einer schnelleren Konvergenz führen. Die Reduzierung der notwendigen Iterationen, bis ein definiertes Misfitniveau erreicht wird, spielt in der Wellenforminversion eine entscheidende Rolle, da die Rechenzeit zurzeit durchaus eine kritische Größe darstellt.

4.5.3.1. Herleitung der konjugierten Gradienten

Für die Herleitung der konjugierten Gradienten wird an dieser Stelle bereits eine Verbindung zum konkreten Misfit hergestellt. Zu beachten ist jedoch, dass zunächst angenommen wird, die quadratische Form der untersuchten Funktion $E(\vec{x})$ sei exakt, d.h. ohne Näherung quadratisch im Argument \vec{x} . Das gewonnene Ergebnis wird anschließend in Kapitel 4.5.3.2 auf nichtlineare Funktionen verallgemeinert.

Die für die Wellenforminversion grundlegende Taylorentwicklung der Misfitfunktion aus Kapitel 4.4

$$S(\vec{m}) = S(\vec{m}_i + \Delta\vec{m}_i) \approx S(\vec{m}_i) + \nabla_{\vec{m}} S(\vec{m}_i)^T \Delta\vec{m}_i + \frac{1}{2} \Delta\vec{m}_i^T H_{S,\vec{m}}(\vec{m}_i) \Delta\vec{m}_i + \mathcal{O}(\Delta\vec{m}_i^3) \quad (4.18)$$

ist äquivalent zu der allgemeinen quadratischen Form

$$E(\vec{x}) = -\langle \vec{b}, \vec{x} \rangle + \frac{1}{2} \langle A\vec{x}, \vec{x} \rangle + C \quad (4.19)$$

$$\text{mit } C = S(\vec{m}_i), \quad \vec{b} = -\nabla_{\vec{m}} S(\vec{m}_i), \quad \vec{x} = \Delta \vec{m}_i, \quad A = H_{S, \vec{m}}(\vec{m}_i)$$

und der Eigenschaft des Standardskalarprodukts $\langle A\vec{x}, \vec{y} \rangle = (A\vec{x})^T \vec{y} = \vec{x}^T A^T \vec{y}$. An der Nullstelle der Ableitung von $S(\vec{m})$ nach $\Delta \vec{m}$ (vgl. Gleichung 4.13) ist zu erkennen, dass das Minimierungsproblem $\arg \min_{\Delta \vec{m}} S(\vec{m})$ identisch zur Lösung des linearen Gleichungssystems $H_{S, \vec{m}}(\vec{m}_i) \Delta \vec{m}_i = -\nabla_{\vec{m}} S(\vec{m}_i)$ ist (bzw. $\arg \min_{\vec{x}} E(\vec{x}) \Leftrightarrow A\vec{x} = \vec{b}$). Dies gilt für symmetrische und positiv definite Matrizen $H_{S, \vec{m}}(\vec{m}_i)$ bzw. A , was von der Hesse-Matrix in einem Minimum erfüllt wird.

Aufgrund dieses Zusammenhangs erlangte das Verfahren der konjugierten Gradienten große Bedeutung als effektives Iterationsverfahren zur Lösung großer linearer Gleichungssysteme. Es geht auf die Arbeit von [Hestenes and Stiefel \(1952\)](#) zurück.

Zur Minimierung einer durch eine quadratischen Form angenäherte Funktion, wie es in dieser Arbeit für den Misfit $S(\vec{m})$ der Fall ist, existieren verschiedene Varianten der konjugierten Gradienten, die hierfür verbesserte Konvergenzverhalten einführen. ([Shewchuk et al., 1994](#))

Eigenschaften des Gradientenverfahrens (Darstellung nach [Neumann \(2013\)](#))

Man betrachte aus Gründen der Übersichtlichkeit die abstrahierte quadratische Form

$$E(\vec{x}) = -\langle \vec{b}, \vec{x} \rangle + \frac{1}{2} \langle A\vec{x}, \vec{x} \rangle + C = -\vec{b}^T \vec{x} + \frac{1}{2} \vec{x}^T A \vec{x} \quad (4.20)$$

welche mit $C = 0$, wie oben gezeigt, äquivalent zu dem entwickelten Misfit ist. Der Gradient in Komponenten ist

$$\begin{aligned} (\nabla_{\vec{x}} E(\vec{x}))_i &= \partial_i E(\vec{x}_j) = \partial_i (-b_j x_j + \frac{1}{2} x_k A_{kl} x_l) = -b_j \delta_{ij} + \frac{1}{2} A_{kl} \partial_i (x_k x_l) \\ &= -b_i + \frac{1}{2} A_{kl} \delta_{ik} x_l + \frac{1}{2} A_{kl} \delta_{il} x_k = -b_i + \frac{1}{2} A_{il} x_l + \frac{1}{2} x_k A_{ki} \\ &= -b_i + \frac{1}{2} A_{il} x_l + \frac{1}{2} A_{ik} x_k = -b_i + \frac{1}{2} A_{il} x_l + \frac{1}{2} A_{il} x_l \\ &= -b_i + A_{il} x_l \end{aligned}$$

Dabei wurde in Zeile 3 verwendet, dass A symmetrisch ist ($A_{ki} = A_{ik}$) und doppelte Indizes umbenannt werden dürfen ($k \rightarrow l$).

$$\nabla_{\vec{x}} E(\vec{x}) = -\vec{b} + A\vec{x} = -\vec{g} \quad (4.21)$$

Um das Minimum von E zu finden, kann die Nullstelle des Gradienten berechnet werden ($\vec{x}_{\text{Newton}} = A^{-1}\vec{b}$). Dieses Ergebnis ist bereits aus Gleichung 4.13 bekannt. \vec{x}_{Newton} zeigt direkt auf das Minimum (Newton-Richtung), erfordert bei optimaler Wahl der Schrittweite also keine Iterationen. Es muss jedoch die Inverse A^{-1} bestimmt werden. Das Gradientenverfahren umgeht dieses Problem, indem sich dem Minimum iterativ in Richtung $\nabla_{\vec{x}} E(\vec{x})$ genähert wird. Das Schema lautet

$$\vec{x}_{i+1} = \vec{x}_i - \alpha_i \nabla_{\vec{x}} E(\vec{x}_i) = \vec{x}_i + \alpha_i \vec{g}_i \quad (4.22)$$

mit der Schrittweite α_i . \vec{x}_0 wird zu Beginn beliebig gewählt.

Die Bestimmung der Schrittweite α_i stellt ein eigenes Teilproblem dar, dem ein eigenes Kapitel (4.6) gewidmet ist. Für exakt quadratische Zielfunktionen lässt sich die optimale Schrittweite leicht analytisch ausdrücken (*exakte Liniensuche*). Dazu wird das Minimum von $E(\vec{x} + \alpha\vec{g})$ in Richtung $\vec{g} = -\nabla_{\vec{x}} E(\vec{x})$ bezüglich α gesucht.

$$\begin{aligned} E(\vec{x} + \alpha\vec{g}) &= -\vec{b}^T(\vec{x} + \alpha\vec{g}) + \frac{1}{2}(\vec{x} + \alpha\vec{g})^T A(\vec{x} + \alpha\vec{g}) \\ &= \underbrace{-\vec{b}^T\vec{x} + \frac{1}{2}\vec{x}^T A\vec{x}}_{E(\vec{x})} - \vec{b}^T\alpha\vec{g} + \frac{1}{2}\alpha\vec{g}^T A(\vec{x} + \alpha\vec{g}) + \frac{1}{2}\alpha\vec{x}^T A\vec{g} \\ &= E(\vec{x}) + \frac{1}{2}\alpha(\vec{g}^T A\vec{x} + \alpha\vec{g}^T A\vec{g} + \vec{x}^T A\vec{g}) - \alpha\vec{b}^T\vec{g} \\ &= E(\vec{x}) + \frac{1}{2}\alpha^2\vec{g}^T A\vec{g} + \alpha\vec{x}^T A\vec{g} - \alpha\vec{b}^T\vec{g} \\ &= E(\vec{x}) + \frac{1}{2}\alpha^2\vec{g}^T A\vec{g} + \alpha \underbrace{(\vec{x}^T A - \vec{b}^T)}_{(A\vec{x} - \vec{b})^T = -\vec{g}} \vec{g} \\ &= E(\vec{x}) + \frac{1}{2}\alpha^2\vec{g}^T A\vec{g} - \alpha\vec{g}^T\vec{g} \end{aligned}$$

Hier wurde $\vec{g}^T A \vec{x} = g_i A_{ij} x_j = x_j A_{ji} g_i = \vec{x}^T A \vec{g}$ und $(\vec{x}^T A - \vec{b}^T)_i = x_j A_{ji} - b_i = A_{ij} x_j - b_i = (A \vec{x} - \vec{b})_i^T$ verwendet.

$$\rightarrow \frac{\partial E(\vec{x} + \alpha \vec{g})}{\partial \alpha} = \alpha \vec{g}^T A \vec{g} - \vec{g}^T \vec{g} = 0 \quad (4.23)$$

$$\alpha_i = \frac{\vec{g}_i^T \vec{g}_i}{\vec{g}^T A \vec{g}} \quad (4.24)$$

Damit lässt sich zeigen, dass aufeinander folgende Gradienten senkrecht aufeinander stehen.

$$\begin{aligned} \vec{g}_i^T \vec{g}_{i+1} &\stackrel{4.21}{=} -\vec{g}_i^T (\vec{b} - A \vec{x}_{i+1}) && \stackrel{4.22}{=} -\vec{g}_i^T (\vec{b} - A(\vec{x}_i + \alpha_i \vec{g}_i)) \\ &= -\vec{g}_i^T (\vec{b} - A \vec{x}_i - \alpha_i A \vec{g}_i) && = \vec{g}_i^T (A \vec{x}_i - \vec{b}) + \alpha_i A \vec{g}_i^T \vec{g}_i \\ &\stackrel{4.21}{=} -\vec{g}_i^T \vec{g}_i + \alpha_i \vec{g}_i^T A \vec{g}_i && \stackrel{4.24}{=} 0 \quad \Rightarrow \vec{g}_i \perp \vec{g}_{i+1} \end{aligned}$$

Dies ist ein anschaulicher Grund für die mäßige Konvergenzgeschwindigkeit. Analysiert man das vorliegende Verfahren, so stellt man fest, dass in der i -ten Iteration lediglich nach der Schrittweite α_i optimiert wird und die vorhergehenden Abstiegsrichtungen \vec{g}_k und Schrittweiten α_k mit $k < i$ als gegeben angesehen werden.

Eigenschaften der konjugierten Gradienten Der Ansatz zum Auffinden der konjugierten Gradienten besteht darin, die optimalen Schrittlängen α_i gleichzeitig, d.h. im Rückblick auf eine Folge von Iterationen, zu suchen. Dazu betrachte man eine Folge von N Schritten noch unbekannter Richtungen \vec{p}_i gemäß des Abstiegsschemas

$$\vec{x}_{i+1} = \vec{x}_i + \alpha \vec{p}_i \quad (4.25)$$

$$\vec{x} = \vec{x}_0 + \alpha_0 \vec{p}_0 + \alpha_1 \vec{p}_1 + \dots + \alpha_N \vec{p}_N = \vec{x}_0 + \sum_{i=0}^N \alpha_i \vec{p}_i$$

Zu dieser Schrittfolge werden nun die einzelnen optimalen Schrittlängen gesucht.

$$\begin{aligned} E \left(\vec{x}_0 + \sum_{i=0}^N \alpha_i \vec{p}_i \right) &\stackrel{4.20}{=} -\vec{b}^T \left(\vec{x}_0 + \sum_i \alpha_i \vec{p}_i \right) + \frac{1}{2} \left(\vec{x}_0 + \sum_i \alpha_i \vec{p}_i \right)^T A \left(\vec{x}_0 + \sum_i \alpha_i \vec{p}_i \right) \\ &= -\vec{b}^T \vec{x}_0 - \vec{b}^T \sum_i \alpha_i \vec{p}_i + \frac{1}{2} \vec{x}_0^T A \vec{x}_0 + \frac{1}{2} \left(\sum_i \alpha_i \vec{p}_i \right)^T A \left(\sum_i \alpha_i \vec{p}_i \right) + \vec{x}_0^T A \sum_i \alpha_i \vec{p}_i \\ &\stackrel{4.21}{=} -\vec{b}^T \vec{x}_0 + \frac{1}{2} \vec{x}_0^T A \vec{x}_0 + \frac{1}{2} \left(\sum_i \alpha_i \vec{p}_i \right)^T A \left(\sum_i \alpha_i \vec{p}_i \right) - \vec{g}_0^T \left(\sum_i \alpha_i \vec{p}_i \right) \end{aligned}$$

$$\begin{aligned}
\frac{\partial E\left(\vec{x}_0 + \sum_{i=0}^N \alpha_i \vec{p}_i\right)}{\partial \alpha_j} &= -\vec{g}_0^T \left(\sum_i \delta_{ij} \vec{p}_i\right) + \frac{1}{2} \left(\sum_i \delta_{ij} \vec{p}_i\right)^T A \left(\sum_i \alpha_i \vec{p}_i\right) + \frac{1}{2} \left(\sum_i \alpha_i \vec{p}_i\right)^T A \left(\sum_i \delta_{ij} \vec{p}_i\right) \\
&= -\vec{g}_0^T \vec{p}_j + \frac{1}{2} \vec{p}_j^T A \left(\sum_i \alpha_i \vec{p}_i\right) + \frac{1}{2} \left(\sum_i \alpha_i \vec{p}_i\right)^T A \vec{p}_j \\
&= -\vec{g}_0^T \vec{p}_j + \left(\sum_i \alpha_i \vec{p}_i\right)^T A \vec{p}_j = 0
\end{aligned} \tag{4.26}$$

Dies stellt ein Gleichungssystem mit $j = 1, \dots, N$ Gleichungen dar. Es ist leicht zu lösen, wenn $\vec{p}_i^T A \vec{p}_j = 0$ für alle $i \neq j$. Mathematisch sind damit alle Abstiegsrichtungen paarweise zueinander A-orthogonal (A-konjugiert).

Die Bestimmung der optimalen Schrittweiten findet analog zur Ausführung in Gleichung 4.26 statt. Nun, wie bereits aus Gleichung 4.23 bekannt, ausgehend von $E(\vec{x}_i + \alpha_i \vec{p}_i)$

$$\frac{\partial E(\vec{x}_i + \alpha_i \vec{p}_i)}{\partial \alpha_i} = -\vec{g}_i^T \vec{p}_i + \alpha_i \vec{p}_i^T A \vec{p}_i = 0 \tag{4.27}$$

$$\alpha_i = \frac{\vec{g}_i^T \vec{p}_i}{\vec{p}_i^T A \vec{p}_i} \tag{4.28}$$

Die A-Orthogonalität der \vec{p} impliziert eine lineare Unabhängigkeit, was bedeutet, dass das Verfahren nach höchstens M Schritten abbricht, wobei M die Dimension des Modells ist ($\vec{p} \in \mathbb{R}^M, \vec{x} \in \mathbb{R}^M$). Anhand von zwei Dimensionen kann bereits gesehen werden, dass damit ein deutlich verbessertes Konvergenzverhalten einhergeht.

Die Bestimmung der Schrittweite α_j in der Iteration j erfordert die Kenntnis der konjugierten Richtung \vec{p}_j . Für deren Berechnung macht man den rekursiven Ansatz

$$\vec{p}_{i+1} = \vec{g}_{i+1} + \beta_i \vec{p}_i \quad \text{mit} \quad \vec{g}_0 = -\nabla_{\vec{x}} E(\vec{x}_0) = \vec{b} - A\vec{x}_0 \tag{4.29}$$

Per Konstruktion basiert jeder konjugierte Gradient \vec{p}_{i+1} auf dem Gradient im derzeitigen Punkt \vec{x}_{i+1} , der um den skalierten konjugierten Gradient aus vorhergehender Iteration $\beta_i \vec{p}_i$ modifiziert wird. Die β_i werden stets so gewählt, dass \vec{p}_{i+1} zu \vec{p}_i A-orthogonal ist.

$$\begin{aligned}
\vec{p}_{i+1}^T A \vec{p}_i &\stackrel{4.29}{=} (\vec{g}_{i+1} + \beta_i \vec{p}_i)^T A \vec{p}_i = 0 \\
\rightarrow \beta_i^{\text{FR}} &= -\frac{\vec{g}_{i+1}^T A \vec{p}_i}{\vec{p}_i^T A \vec{p}_i} \quad (\text{Fletcher-Reeves})
\end{aligned} \tag{4.30}$$

Der Ansatz stellt lediglich die A-Orthogonalität von aufeinander folgenden Richtungen \vec{p}

sicher. Eleganter Weise kann jedoch gezeigt werden, dass das nach Gleichung 4.30 gewählte β_i dazu führt, dass alle \vec{p} paarweise A-konjugiert sind, wie es gefordert wurde.

Zusammenfassend sei das Verfahren der konjugierten Gradienten zur Lösung von linearen Gleichungssystemen $A\vec{x} = \vec{b}$ und Minimierung von quadratischen Funktionen $E(\vec{x})$ in M Dimensionen aufgeführt.

1. Wähle Startpunkt \vec{x}_0 .
2. Berechne reinen Gradienten in \vec{x}_0 als erste Abstiegsrichtung (Gleichung 4.29).

$$\vec{p}_0 = -\vec{g}_0 = -\nabla_{\vec{x}}E(\vec{x}_0) = \vec{b} - A\vec{x}_0$$

3. Iterative Modellverbesserung entlang des konjugierten Gradienten \vec{p}_i mit Schrittweite.

$$\begin{aligned} \vec{x}_{i+1} &\stackrel{4.25}{=} \vec{x}_i + \alpha_i \vec{p}_i && \text{(Abstiegsschema)} \\ \alpha_i &\stackrel{4.28}{=} \frac{\vec{g}_i^T \vec{p}_i}{\vec{p}_i^T A \vec{p}_i} && \text{(Schrittweite)} \\ \vec{g}_{i+1} &\stackrel{4.21}{=} \vec{b} - A\vec{x}_{i+1} && \text{(reiner Gradient)} \\ \vec{p}_{i+1} &\stackrel{4.29}{=} \vec{g}_{i+1} + \beta_i^{\text{FR}} \vec{p}_i && \text{(konjugierter Gradient)} \\ \beta_i^{\text{FR}} &\stackrel{4.30}{=} -\frac{\vec{g}_{i+1}^T A \vec{p}_i}{\vec{p}_i^T A \vec{p}_i} && \text{(Konjugation des Gradienten)} \end{aligned}$$

4.5.3.2. Erweiterung der konjugierten Gradienten auf nichtlineare Funktionen

Das vorhergehend beschriebene Verfahren ist zur Minimierung von Funktionen $E(\vec{x})$ gültig, die exakt quadratisch in \vec{x} sind. Die für die Wellenforminversion relevante Misfit Funktion $S(\vec{m})$ wurde bis zweiter Ordnung in $\Delta\vec{m}$ entwickelt und besitzt daher in der Umgebung des Entwicklungspunktes näherungsweise die quadratische Form (vgl. Gl. 4.19) auf der die konjugierten Gradienten basieren. Für die Anwendung auf allgemeine nichtlineare Zielfunktionen sind zwei Modifikationen notwendig.

Die Suche nach der optimalen Schrittweite α , welche die Zielfunktion $S(\vec{m})$ entlang der Abstiegsrichtung \vec{p} minimiert, kann nicht mehr analytisch durch eine *exakte Liniensuche*

erfolgen. An ihre Stelle tritt eine *inexakte Liniensuche*, die das Minimum von $S(\vec{m})$ entlang \vec{p} annähert. Dieses Thema wird ausführlich in Kapitel 4.6 behandelt. Die zweite Änderung betrifft den Gradienten \vec{g} , für den die analytischen Ausdrücke nicht gelten. Er wird direkt durch den Gradienten $\nabla_{\vec{m}}S(\vec{m})$ ersetzt (beispielsweise als Resultat aus der adjoint state Methode).

Des Weiteren existieren verschiedene Varianten zur Konjugation, die angewendet auf den quadratischen Fall äquivalent sind. Neben der für quadratische Zielfunktionen hergeleiteten Form nach Fletcher-Reeves hat sich die Formel nach Polak-Ribière speziell für den nichtlinearen Fall als besonders praktikabel herausgestellt. Numerische Experimente haben gezeigt, dass die Wahl von β nach Polak-Ribière robuster ist und schnellere Konvergenz aufweist (Nocedal and Wright, 2006).

$$\beta_i^{\text{PR}} = \frac{\nabla_{\vec{m}}S(\vec{m}_{i+1})^T(\nabla_{\vec{m}}S(\vec{m}_{i+1}) - \nabla_{\vec{m}}S(\vec{m}_i))}{\nabla_{\vec{m}}S(\vec{m}_i)^T\nabla_{\vec{m}}S(\vec{m}_i)} \quad (4.31)$$

β_i^{PR} alleine stellt noch nicht sicher, dass \vec{p} eine Abstiegsrichtung ist. Hierzu ist es erforderlichlich $\beta_i^{\text{PR}+}$ nach

$$\beta_i^{\text{PR}+} = \max\{\beta_i^{\text{PR}}, 0\} \quad (4.32)$$

zu wählen. Falls $\beta_i^{\text{PR}} < 0$ wird $\beta_i^{\text{PR}+}$ damit auf 0 gesetzt, was einem Neustart des konjugierten Gradientenverfahrens entspricht, d.h. in dieser Iteration wird der reine Gradient als Abstiegsrichtung verwendet.

Generell ist das Auffinden des globalen Minimums einer allgemeinen Zielfunktionen eine größere Herausforderung. Im Gegensatz zur quadratischen Form können lokale Minima auftreten, wobei das Verfahren nicht sicherstellen kann, dass es ins globale Minimum konvergiert.

Damit lautet das Verfahren der konjugierten Gradienten für nichtlineare Zielfunktionen (hier in der Notation des Modells \vec{m} und der Zielfunktion des Misfits $S(\vec{m})$):

1. Wähle Startmodell \vec{m}_0 .
2. Berechne reinen Gradienten in \vec{m}_0 als erste Abstiegsrichtung.

$$\vec{p}_0 = -\nabla_{\vec{m}}S(\vec{m}_0)$$

3. Iterative Modellverbesserung entlang des konjugierten Gradienten \vec{p}_i mit Schrittweite.

$$\begin{aligned} \vec{m}_{i+1} &\stackrel{4.25}{=} \vec{m}_i + \alpha_i \vec{p}_i && \text{(Abstiegsschema)} \\ \alpha_i & && \text{(inexakte Liniensuche, s. Kapitel 4.6)} \\ \nabla_{\vec{m}} S(\vec{m}_{i+1}) & && \text{(adjoint state Methode s. Kapitel 4.5.2)} \\ \vec{p}_{i+1} &\stackrel{4.29}{=} -\nabla_{\vec{m}} S(\vec{m}_{i+1}) + \beta_i^{\text{PR+}} \vec{p}_i && \text{(konjugierter Gradient)} \\ \beta_i^{\text{PR+}} &\stackrel{4.32}{=} \max\{\beta_i^{\text{PR}}, 0\} && \text{(Konjugation des Gradienten)} \end{aligned}$$

(Shewchuk et al., 1994)

4.5.4. Prädiktionierung des Gradienten

Da der Gradient $\nabla_{\vec{m}} S(\vec{m})$ der Zielfunktion S bezüglich der Modellparameter \vec{m} auch als Maß der Sensitivität gesehen werden kann mit dem der Misfit S auf die Änderung von einzelnen Modellparameter reagiert, ist es direkt ersichtlich, dass der Wert dieses Gradienten S in den direkten Umgebungen der Quelllokationen höher ist als im übrigen Modell. Die Energie der Wellen ist an dieser Stelle auf einen sehr kleinen Raum konzentriert, wodurch hier ein Großteil der Energie durch Modellperturbationen gestört wird. Diese Eigenschaft äußert sich in Inversionsergebnissen durch sogenannte Quellartefakte. Diese haben zwar wie beschrieben einen physikalisch begründeten Ursprung, doch sind in der Anwendung die Quellen oftmals durch hohe Werte im Modell sichtbar, was geologisch nicht nachvollziehbar ist.

Eine Gegenmaßnahme ist die Modifikation des Gradienten $\nabla_{\vec{m}} S(\vec{m})$ bzw. allgemein der Abstiegsrichtung \vec{p}_i durch eine Prädiktionierung. Ein empirischer Ansatz zur Konstruktion dieser Prädiktionierungsmatrix P ist die Verwendung eines Quelltapers. Hier wird der Gradient im Bereich der Quellpositionen sanft auf 0 gesetzt. Für marine Datensätze ist der Quelltaper meist obsolet, da die Quellpositionen im Meer liegen. Die Materialparameter des Wassers sind meist bekannt und werden bereits im Startmodell angesetzt. Um Modellupdates hier zu verhindern kann ein Taper verwendet werden, der den Gradient auf 0 setzt, womit die Quellpositionen eingeschlossen sind.

Eine physikalisch motivierte Methode der Prädiktionierung nach [Shin et al. \(2001\)](#) ist als weitere konvergenzverbessernde Maßnahme des Gradientenverfahrens an der Grenze zu den Quasi Newton Verfahren zu sehen, da sie Informationen der Hessematrix verwendet. Sie basiert auf der Annahme, dass die Diagonalelemente einer approximierten Hessematrix H_a diese approximierte Hessematrix H_a in der Nähe des globalen Minimums des Misfits S gut nähert. Damit stellt die inverse Diagonale der approximierten Hessematrix ein geeigneter Prädiktionierer P dar. Es kann gezeigt werden, dass P die Skalierung des Gradienten „normalisiert“, d.h. Quellartefakte werden reduziert. Weiterhin sind die Diagonalelemente stark vom geometrischen Ausbreitungsverlust beeinflusst, was hilft den Gradienten besser zu skalieren und das Modellupdate in großem Quelle-Empfänger-Abstand zu verstärken.

Die Herleitung der approximierten Hessematrix H_a kann in wenigen Zeilen nachvollzogen werden. Die Hessematrix enthält die zweiten partiellen Ableitungen des Misfits S nach den Modellparametern \vec{m} ($H_{S,\vec{m}}(\vec{m}) = \frac{\partial^2 S}{\partial m_i \partial m_j}$). Mit der Definition des Misfits aus Gleichung 4.2 gilt für $H_{S,\vec{m}}(\vec{m})$ (\vec{d} : observierte Daten, \vec{u} : synthetische Daten)

$$H_{S,\vec{m}}(\vec{m}) = \frac{\partial^2}{\partial m_i \partial m_j} \left(\frac{1}{2} \sum_{l=1}^N (d_l - u_l)^2 \right) = \frac{\partial}{\partial m_i} \left(- \sum_{l=1}^N (d_l - u_l) \frac{\partial u_l}{\partial m_j} \right) \quad (4.33)$$

$$= \underbrace{\sum_{l=1}^N \frac{\partial u_l}{\partial m_i} \frac{\partial u_l}{\partial m_j}}_{H_a} - \underbrace{\sum_{l=1}^N (d_l - u_l) \frac{\partial^2 u_l}{\partial m_i \partial m_j}}_R \quad (4.34)$$

Die Inverse einer Diagonalmatrix berechnet sich durch die Inversion der einzelnen Diagonalelemente. Randbereiche im Modell mit geringer Strahlabdeckung führen zu sehr kleinen Einträgen, die in H_a^{-1} dann sehr groß werden und das Modellupdate hier extrem verstärken. Darum schlägt [Shin et al. \(2001\)](#) vor die Inversion von H_a durch Verwendung eines Dämpfungsfaktors (waterlevel) χ zu stabilisieren. Schließlich folgt für das Iterationsschema mit prädiktioniertem Gradienten (α_i : Schrittweite):

$$\vec{m}_{i+1} = \vec{m}_i + \alpha_i \vec{p}_i = \vec{m}_i + \alpha_i \underbrace{(\text{diag}(H_{a,i}) + \chi \mathbf{1})^{-1}}_{P_i} \nabla_{\vec{m}} S(\vec{m}_i) \quad (4.35)$$

Stehen mehrere Schüsse zur Verfügung, so wird der Gradient und der Prädiktionierer für jeden Schuss getrennt berechnet. Es besteht dann die Möglichkeit Gradient und Prädiktionierer für jeden Schuss separat aufzusummieren und die Summe von P auf die Summe des Gradienten anzuwenden (Prädiktionierung per Summe) oder aber den Prä-

konditionierer direkt auf den Gradienten jedes Schusses anzuwenden (Präkonditionierung per Schuss) (siehe dazu das Ablaufdiagramm in Abbildung 4.8).

4.5.5. Newton und Quasi Newton Verfahren

Werden die inverse Hesse-Matrix H^{-1} und der Gradient für die Modellverbesserung verwendet, spricht man von Newton Verfahren. Zu deren Anwendung müssen alle $\frac{n(n+1)}{2}$ partiellen Ableitungen 2. Ordnung berechnet, sowie die Inverse von H bestimmt werden, was in der Praxis der FWI aufgrund der Vielzahl an Modellparameter zur Zeit kaum realisierbar ist. Vorteilhaft ist die lokal quadratische Konvergenzgeschwindigkeit, d.h. die Anzahl der genauen Dezimalstellen verdoppelt sich bei jeder Iteration.

Neben der reinen Gradientenmethode stellen die Quasi Newton Verfahren eine Alternative dar, welche auf dem Konzept des Newton Verfahrens basieren. Hierbei wird die Hessematrix (direkte Quasi Newton Verfahren) oder unmittelbar die inverse Hessematrix (inverse Quasi Newton Verfahren) approximiert. Trotz deutlich geringerem Aufwand liegt meist eine lokal superlineare Konvergenz vor.

Idee der Quasi Newton Verfahren ist die explizite Berechnung der Hesse-Matrix $H_{S,\vec{m}}(\vec{m}_i)$ durch ein iteratives Verfahren zu ersetzen, bei dem Matrizen M_i berechnet werden für die gilt, dass der Übergang von M_i zu M_{i+1} einfach ist und M_i $H_{S,\vec{m}}(\vec{m}_i)$ approximiert. Weitere Forderungen sind: M_i positiv definit (Extremum von S ist Minimum) und symmetrisch.

Zur Konstruktion des Quasi Newton Verfahrens wird aus Gleichung 4.13 ein Ausdruck für den Gradienten von S im Modell \vec{m}_i und im nächsten Iterationsschritt \vec{m}_{i+1} gewonnen. Dabei gilt $\Delta\vec{m}_i = \vec{m} - \vec{m}_i$ bzw. $\Delta\vec{m}_{i+1} = \vec{m} - \vec{m}_{i+1}$.

$$\nabla_{\vec{m}}S(\vec{m}_i) = -H_{S,\vec{m}}(\vec{m}_i)(\vec{m} - \vec{m}_i) \quad (4.36)$$

$$\nabla_{\vec{m}}S(\vec{m}_{i+1}) = -H_{S,\vec{m}}(\vec{m}_{i+1})(\vec{m} - \vec{m}_{i+1}) \quad (4.37)$$

Die Differenz der aufeinander folgenden Gradienten ergibt:

$$\nabla_{\vec{m}}S(\vec{m}_{i+1}) - \nabla_{\vec{m}}S(\vec{m}_i) = -H_{S,\vec{m}}(\vec{m}_{i+1})(\vec{m} - \vec{m}_{i+1}) + H_{S,\vec{m}}(\vec{m}_i)(\vec{m} - \vec{m}_i) \quad (4.38)$$

Es wird nun angenommen, dass die Hessematrix an der Stelle m_i und m_{i+1} ungefähr

gleich sind. Diese Näherung zeichnet das Quasi Newton Verfahren gegenüber dem Newton Verfahren aus und führt auf die Sekanten oder Quasi Newton Gleichung (QNG).

$$\nabla_{\vec{m}} S(\vec{m}_{i+1}) - \nabla_{\vec{m}} S(\vec{m}_i) \approx H_{S,\vec{m}}(\vec{m}_{i+1})(\vec{m}_{i+1} - \vec{m}_i) \quad (4.39)$$

Die Bezeichnung als Sekantengleichung kann über eine Analogie zum Differenzenquotienten eingesehen werden. Die zweite Ableitung wird aus einer Sekante an die erste Ableitung bestimmt.

$$H_{S,\vec{m}}(\vec{m}_{i+1}) \approx \frac{\nabla_{\vec{m}} S(\vec{m}_{i+1}) - \nabla_{\vec{m}} S(\vec{m}_i)}{\vec{m}_{i+1} - \vec{m}_i} \quad (4.40)$$

Die Lösung der QNG ist nicht eindeutig und positive Definitheit ist nicht sichergestellt, was erfüllt sein muss, damit das Extremum der Zielfunktion ein Minimum darstellt.

Es existieren verschiedene Methoden Lösungen der QNG zu bestimmen. Das Verfahren nach Davidon-Fletcher-Powell (DFP, 1959), war die erste verallgemeinerte Quasi Newton Lösung für mehrdimensionale Probleme. Sie symmetrische Rang-1-Methode (SR1) stellt Eindeutigkeit aber keine positive Definitheit sicher. Die Broyden-Fletcher-Goldfarb-Shanno-Methode (BFGS) (s. Kapitel 4.5.7) beinhaltet die SR1, führt auf positiv definite Matrizen und hat DFP-Verfahren weitgehend ersetzt. Die Erweiterung L-BFGS (L: low memory) hat einen reduzierten Speicherbedarf und eignet sich besonders für die Anwendung in der FWI.

4.5.6. Interpretation der Newton Richtung

Im Iterationsschema aus Gleichung 4.14 wird der negative Gradient $-\nabla_{\vec{m}} S(\vec{m}_i)$, welcher in Richtung des steilsten Abstiegs der Misfitfunktion zeigt, sowie die Hessematrix $H_{S,\vec{m}}(\vec{m}_i)$ verwendet, die alle partiellen Ableitungen 2. Ordnung der Misfitfunktion S nach allen Modellparametern \vec{m} an der Stelle \vec{m}_i enthält. Diese beschreibt die Krümmungseigenschaften von S . Nachfolgend soll anhand eines Beispiels gezeigt werden, dass die Verwendung der Hessematrix im Vergleich zum reinen Gradienten zu schnellerer Konvergenz führt.

Betrachtet wird eine quadratische Misfitfunktion S in zwei Dimensionen, d.h. das Modell ist zweidimensional. Eine solche Funktion, allerdings mit n-dimensionalem Modell, kann man sich als lokale Approximation an S vorstellen, welche in Kapitel 4.4 ebenfalls bis zweiter Ordnung entwickelt wurde. Hier wird ein Ellipsoid $S(\vec{m}) = -\sqrt{\left(1 - \frac{m_1^2}{a^2} - \frac{m_2^2}{b^2}\right)} c^2$

verwendet, das für $a = 5, b = 2, c = 1$ in Abbildung 4.4 dargestellt ist.

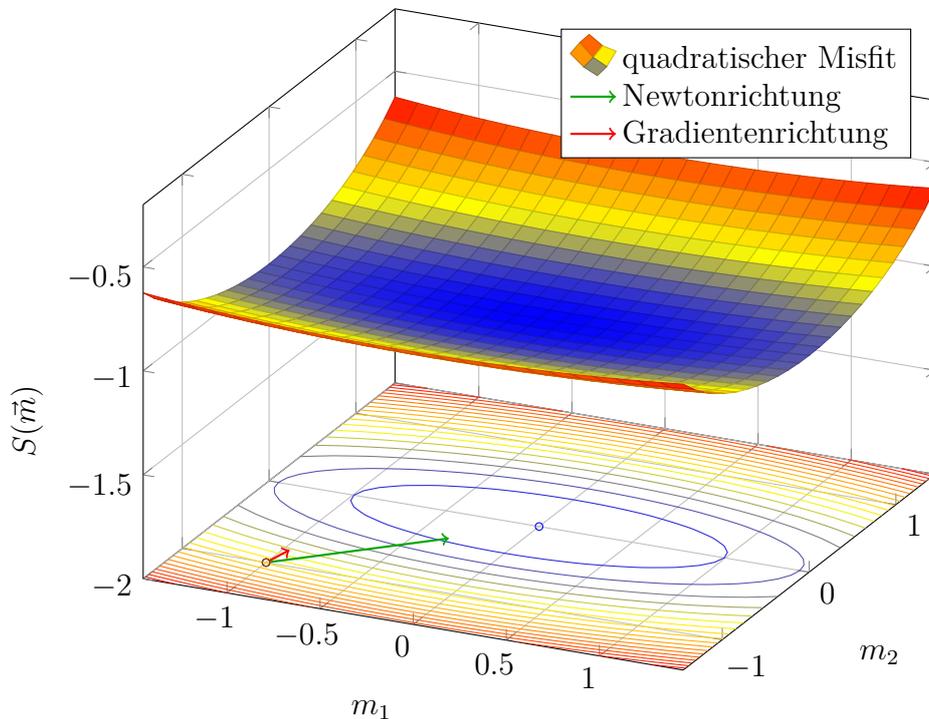


Abbildung 4.4.: Vergleich zwischen Gradienten- und Newtonrichtung am Beispiel einer quadratischen (ellipsoidalen) Misfitfunktion in zwei Modellparameter m_1 und m_2 , eingezeichnet in die Projektionsfläche der Höhenlinien. Die Newton Richtung (grüner Pfeil) weist in Richtung des Minimums, während dies vom negativen Gradient (roter Pfeil) verfehlt wird.

Wird die Modellverbesserung über den Gradienten (Gradienten oder *steepest descent* Methode) erzielt (roter Pfeil), so kann das Minimum nicht in einem Schritt erreicht werden. Wird hingegen die Newton Richtung (grüner Pfeil) mit den zusätzlichen Informationen der Hessematrix verwendet, so kann das Minimum bei geeigneter Wahl der Schrittweite in nur einer Iteration erreicht werden.

4.5.7. BFGS-Verfahren

Ziel ist es eine iterative Rechenvorschrift für die Ermittlung der approximierten Hessematrix M auf Grundlage der QNG (Gl. 4.39) zu finden. Dies wird durch zwei Modifikationen erreicht, welche dafür sorgen, dass M positiv definit und eindeutig ist.

Für eine kompakte Schreibweise definiert man:

$$\vec{s}_i := \vec{m}_{i+1} - \vec{m}_i \quad (4.41)$$

$$\vec{y}_i := \nabla_{\vec{m}} \mathcal{S}(\vec{m}_{i+1}) - \nabla_{\vec{m}} \mathcal{S}(\vec{m}_i) \quad (4.42)$$

Womit sich die QNG (Gl. 4.39) schreibt als:

$$\vec{y}_i = M_{i+1} \vec{s}_i \quad (4.43)$$

Die erste Modifikation lautet:

$$\tilde{M}_i := M_i - \frac{(M_i \vec{s}_i)(M_i \vec{s}_i)^T}{\vec{s}_i^T M_i \vec{s}_i} \quad (4.44)$$

Damit ist \tilde{M}_i symmetrisch und zumindest positiv semidefinit, wenn M_i bereits symmetrisch und positiv definit ist. Als Startmatrix M_0 muss daher eine positiv definite symmetrische Matrix gewählt werden, z.B. eine skalierte Einheitsmatrix $A_0 = \gamma \mathbf{1}$. Mit dieser Modifikation allein erfüllt \tilde{M}_i die QNG noch nicht, denn es gilt:

$$\tilde{M}_i \vec{s}_i = M_i \vec{s}_i - \frac{(M_i \vec{s}_i)(M_i \vec{s}_i)^T}{\vec{s}_i^T M_i \vec{s}_i} \vec{s}_i \quad (4.45)$$

$$= \frac{1}{\vec{s}_i^T M_i \vec{s}_i} \left[(\vec{s}_i^T M_i \vec{s}_i) M_i \vec{s}_i - (M_i \vec{s}_i)(M_i \vec{s}_i)^T \vec{s}_i \right] \quad (4.46)$$

$$= \frac{1}{\vec{s}_i^T M_i \vec{s}_i} \left[(\vec{s}_i^T M_i \vec{s}_i) M_i \vec{s}_i - M_i \vec{s}_i \vec{s}_i^T M_i^T \vec{s}_i \right] = 0 \quad (4.47)$$

Hierbei wurde $(M_i \vec{s}_i)^T = \vec{s}_i^T M_i^T$ und $M_i^T = M_i$, da M_i symmetrisch, verwendet.

Der Rang der Modifikation ist durch den Zähler gegeben. Er hat die Form $(\vec{a}\vec{a}^T)_{ij}$, wodurch sich schreiben lässt:

$$(\vec{a}\vec{a}^T)_{ij} = a_i a_j = \begin{pmatrix} a_1 a_1 & \dots & a_1 a_n \\ a_2 a_1 & \dots & a_2 a_n \\ \vdots & & \vdots \\ a_n a_1 & \dots & a_n a_n \end{pmatrix} = \begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix}, \dots, a_n \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \quad (4.48)$$

Alle Spaltenvektoren sind Linearkombinationen von \vec{s} , weshalb gilt

$$\text{rang} \left((M_i \vec{s}_i) (M_i \vec{s}_i)^T \right) = 1 \quad (4.49)$$

Damit bezeichnet man 4.44 als symmetrische Rang-1-Modifikation.

Durch eine zweite Rang-1-Modifikation wird die positiv definite Matrix erhalten und dafür gesorgt, dass die QNG eindeutig gelöst wird.

$$M_{i+1} = \tilde{M}_i + uZZ^T \quad \text{soll die QNG lösen} \quad (4.50)$$

$$M_{i+1}\vec{s}_i = \underbrace{\tilde{M}_i\vec{s}_i}_{=0} + uZZ^T\vec{s}_i = \vec{y}_i \quad (4.51)$$

$$\rightarrow \text{erfüllt für: } Z = \vec{y}_i \quad \text{und} \quad u = \frac{1}{\vec{y}_i^T \vec{s}_i} \quad (4.52)$$

Schließlich folgt die BFGS Iterationsformel.

$$M_{i+1} = M_i - \frac{(M_i\vec{s}_i)(M_i\vec{s}_i)^T}{\vec{s}_i^T M_i \vec{s}_i} + \frac{\vec{y}_i \vec{y}_i^T}{\vec{y}_i^T \vec{s}_i} \quad (4.53)$$

Die Summe zweier Rang-1 Matrizen ist in der Regel vom Rang 2, weshalb man von einer Rang-2-Modifikation spricht.

4.6. Bestimmung der Schrittweite - Liniensuche

Wie in obigem Beispiel zu sehen, wird die Qualität eines Optimierungsverfahrens neben der Abstiegsrichtung \vec{p}_i , welche im Kontext der FWI der Modellverbesserung $\Delta\vec{m}_i$ entspricht, von der geeigneten Wahl einer Schrittlänge α_i bestimmt.

$$\vec{m}_{i+1} = \vec{m}_i + \alpha_i \Delta\vec{m}_i = \vec{m}_i + \alpha_i \vec{p}_i \quad \text{mit } \alpha_i \in \mathbb{R}_+ \quad (4.54)$$

Die Liniensuche hat das Ziel eine geeignete Schrittweite α_i für eine Modellverbesserung entlang der gewählten Abstiegsrichtung \vec{p}_i für jede Iteration i zu finden. Für die folgenden Betrachtungen ist es hilfreich den Misfitverlauf für das fest gewählte \vec{p}_i als Funktion von α_i mit $F(\alpha_i) = S(\vec{m}_i + \alpha_i \vec{p}_i)$ zu beschreiben. Ein prinzipieller Verlauf von $F(\alpha_i)$ mit seinen charakteristischen Stellen, wie lokale Minima und globales Minimum, ist in Abbildung 4.3 dargestellt. Hilfreich ist der Bezug zu Abbildung 4.4 oder 4.2, wo ein beispielhafter Misfit über zwei Modellparameter dargestellt ist. Die Kurve von $F(\alpha_i)$ ergibt sich als (senkrechter) Schnitt in Richtung des Abstiegs.

Für die Liniensuche stehen zwei Strategien zur Wahl. Die *exakte Liniensuche* ist theoretischer Natur und hat das Ziel das globale Minimum von $F(\alpha_i)$ durch das Lösen von $\frac{\partial F}{\partial \alpha} = 0$ zu finden. In den meisten Anwendungen ist die Berechnung der exakten Schrittweite jedoch zu aufwendig. Eine praktikable Alternative bietet die *inexakte Liniensuche*. Sie besteht aus zwei Phasen. Zuerst wird ein akzeptables Schrittweitenintervall bestimmt. Dies erfolgt durch die Berechnung des Misfits für Testschrittweiten bis dieser spezifische Anforderungen erfüllt. Anschließend wird durch Interpolation innerhalb dieses Intervalls eine optimale Schrittweite gesucht.

Nachfolgend wird die Armijo und Powell Wolfe Regel vorgestellt. Sie stellen eine von vielen Möglichkeiten dar, um akzeptable Schrittweiten zu finden. (Nocedal and Wright, 2006)

4.6.1. Armijo Regel

Die Einhaltung der Armijo Regel soll einen hinreichenden Abstieg garantieren. Für ein fest gewähltes $c_1 \in (0, 1)$ ist $\alpha_i \in (0, 1)$ so zu wählen, dass gilt: (Nocedal and Wright, 2006)

$$S(\vec{m}_i + \alpha_i \vec{p}_i) \leq S(\vec{m}) + c_1 \alpha_i \nabla S(\vec{m})^T \vec{p}_i \quad (4.55)$$

$$\Leftrightarrow F(\alpha_i) \leq F(0) + c_1 \alpha_i F'(0) \quad (4.56)$$

Hier lohnt ein Blick auf das Beispiel aus Kapitel 4.5.6 mit dem Gradienten als Abstiegsrichtung \vec{p} (roter Pfeil in Abbildung 4.4). $F(\alpha_i)$ ergibt den Misfit, wenn ein Schritt der Länge α_i in Richtung \vec{p}_i vollzogen wird (Abbildung 4.3). $F'(\alpha_i)$ ist die Wertänderung von $F(\alpha)$ in α_i . Da c_1 und α_i zwischen 0 und 1 liegt, verläuft $c_1 \alpha_i F'(0)$ flacher als $F'(0)$. Die Armijo Regel besagt, dass α_i so gewählt werden muss, damit der Funktionswert unterhalb der abgeflachten Tangente liegt, womit sichergestellt wird, dass bei großen Schrittweiten auch hohe Misfitreduzierungen erzielt werden (vgl. Abbildung 4.5).

4.6.2. Powell Wolfe Regeln

Zur Verhinderung sehr kleiner (ineffizienter) Schrittweiten wurden die (strenge) Powell Wolfe Regeln eingeführt. Sie wiederholen die Armijo Regel (jedoch mit $c_1 \in (0, \frac{1}{2})$) und

sind damit vollständig. (Nocedal and Wright, 2006)

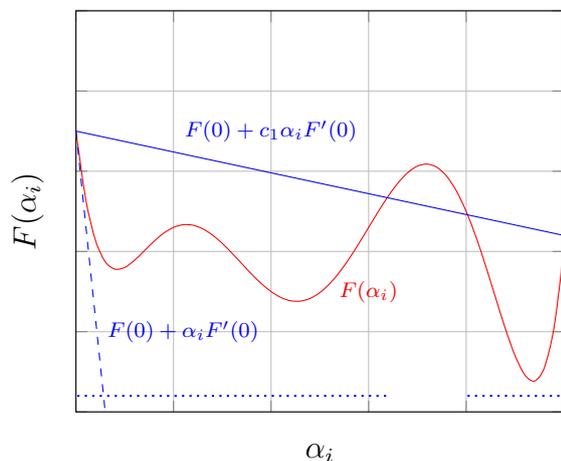


Abbildung 4.5.: Veranschaulichung der Armijo Regel, die ein hinreichender Abstieg sicherstellt. Der Bereich akzeptabler Schrittweiten ist blau gepunktet dargestellt.

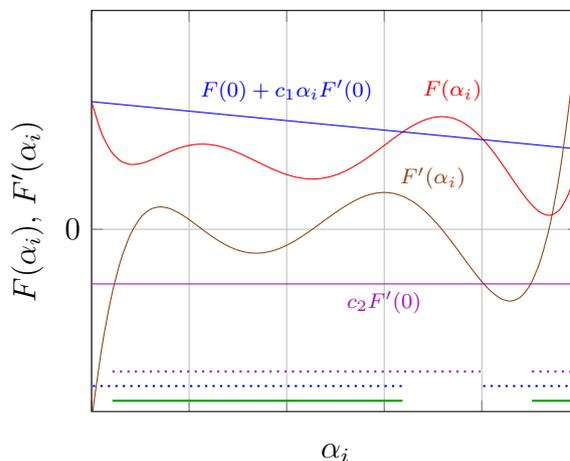


Abbildung 4.6.: Dargestellt ist der Misfitverlauf und dessen Ableitung, auf die sich die Krümmungsbedingung der Powell Wolfe Regel bezieht, um ineffiziente Schrittweiten zu verhindern. Das grüne Schrittweitenintervall genügt beiden Powell Wolfe Regeln.

4.6.2.1. Powell Wolfe Regel

Für ein vorgegebenes $c_1 \in (0, \frac{1}{2})$ und $c_2 \in [c_1, 1)$ ist $\alpha_i > 0$ zu wählen, damit gilt:

$$\nabla S(\vec{m} + \alpha_i \vec{p}_i)^T \vec{p}_i \geq c_2 \nabla S(\vec{m})^T \vec{p}_i \quad (4.57)$$

$$\Leftrightarrow F'(\alpha_i) \geq c_2 F'(0) \quad (4.58)$$

Die Krümmungsbedingung besagt, dass die Steigung von F in α_i größer sein muss als die um c_2 abgeflachte Steigung in $\alpha_i = 0$ (s. Abbildung 4.6). Dies ist sinnvoll, denn wenn die Steigung sehr klein wird, ist dies ein Hinweis, dass durch leichtes Vergrößern der Schrittweite eine signifikante Verkleinerung des Misfits erzielt werden kann.

4.6.2.2. Strenge Powell Wolfe Regel

Für ein vorgegebenes $c_1 \in (0, \frac{1}{2})$ und $c_2 \in [c_1, 1)$ ist $\alpha_i > 0$ zu wählen, damit gilt:

$$|\nabla S(\vec{m} + \alpha_i \vec{p}_i)^T \vec{p}_i| \geq -c_2 \nabla S(\vec{m})^T \vec{p}_i \quad (4.59)$$

$$\Leftrightarrow |F'(\alpha_i)| \geq -c_2 F'(0) \quad (4.60)$$

Hier wird die Wahl von α_i zusätzlich nach oben bis zu einem maximalen Anstieg von F' eingeschränkt.

Die Wolfe Bedingungen eignen sich besonders in Kombination mit dem BFGS-Verfahren (s. Kapitel 4.5.7), das die Hessematrix H_i approximiert. Denn die Einhaltung der 2. Wolfe Bedingung impliziert, dass H_{i+1} positiv definit ist, wenn es H_i ist, was für das BFGS Update notwendig ist.

Für Newton und Quasi Newton Verfahren schlagen [Nocedal and Wright \(2006\)](#) $c_1 = 10^{-4}$ und $c_2 = 0.9$ vor.

4.6.3. Parabolische Interpolation

Die parabolische Interpolation stellt eine einfache Realisierung für die zweite Phase der inexakten Liniensuchen dar. Hierbei wird mit drei Testschrittweiten und deren Misfits eine Parabel interpoliert. Als erster Stützpunkt kann eine Schrittweite mit Wert 0 dienen. Damit kann eine Berechnung des Misfits und Fortwärtsmodellierung eingespart werden, da der Misfit hierzu bereits bekannt ist. Die finale Schrittweite $\alpha_{i,\min}$ ergibt sich aus dem Minimum der Parabel (s. Abbildung 4.7).

An dieser Stelle entstehen viele Probleme auf die es keine optimalen Lösungen gibt. Es muss dann in der jeweiligen Anwendung abgewogen werden, wie aufwendig die Berechnung des Misfits für Testschrittweiten ist. Denn davon hängt ab, ob es sich lohnt die parabolische Schrittweitensuche selbst iterativ durchzuführen, um einem Minimum näher zu kommen oder ob weitere Modellverbesserungen erst nach einer Neuberechnung der Abstiegsrichtung erfolgen soll. Auch kann ein berechnetes $\alpha_{i,\min}$ auf die Powell Wolfe Bedingungen überprüft werden. Auf jeden Fall ist eine Neuberechnung der Schrittweite

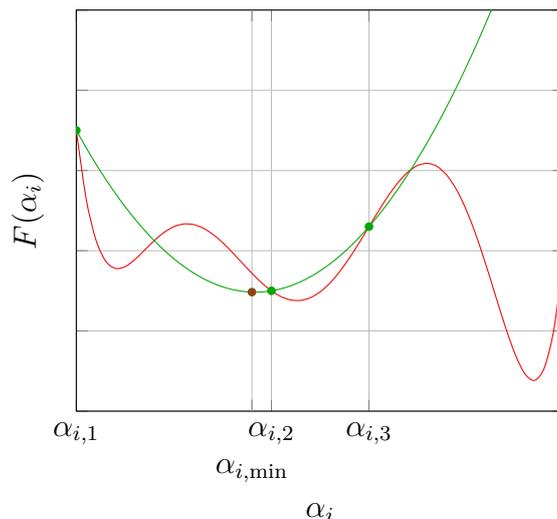


Abbildung 4.7.: Bestimmung einer effizienten Schrittweite $\alpha_{i,\min}$ mithilfe einer parabolischen Interpolation von drei Testschrittweiten $\alpha_{i,1}$, $\alpha_{i,2}$ und $\alpha_{i,3}$.

notwendig, wenn die Parabel nicht nach unten geöffnet ist.

4.7. Inversion des Quellsignals

Die exakte Kenntnis des Quellsignals, das zur Akquisition der Felddaten verwendet wurde, ist fundamental für korrekte Inversionsergebnisse. Denn dieses Signal wird bei der Vorwärtsrechnung zur Erzeugung der synthetischen Seismogramme verwendet. Der Inversionsalgorithmus wird versuchen jedes Datenresiduum, egal ob die Ursache in einem falschen Modell oder einer falschen Signalform liegt, in eine Modellverbesserung zu interpretieren. Prinzipiell wurde für jede der 61 Schussektionen das gleiche Quellsignal verwendet, doch können technische Toleranzen bei der praktischen Durchführung der Schüsse zu leicht unterschiedlichen Signalformen führen. Aus diesem Grund ist es sinnvoll jede Schussektion nach dem Quellsignal zu invertieren.

Zur Auffindung eines Inversionsverfahrens betrachte man die Wellengleichung im Frequenzraum.

$$T(\omega)\vec{u}(\omega) = \vec{f}(\omega) \quad (4.61)$$

$\vec{u}(\omega) \in \mathbb{C}^N$ bezeichnet das fouriertransformierte komplexwertige Wellenfeld. $T(\omega) \in \mathbb{C}^{N \times N}$ wird als Impedanzmatrix bezeichnet. $\vec{f}(\omega) \in \mathbb{C}^N$ ist der Quellterm. (Pratt, 1999)

Zunächst wird der Minimierungsparameter $\kappa(\omega) \in \mathbb{C}$ eingeführt, dessen Bedeutung nicht unmittelbar ersichtlich ist. Er beschreibt über

$$\vec{f}(\omega) = \kappa(\omega) \vec{f}(\omega) \quad (4.62)$$

den Unterschied zwischen zwei Quellsignalen. Diese Signale erzeugen die Wellenfelder $\vec{u}(\omega)$ und $\vec{u}(\omega)$, was sich durch die Wellengleichung ausdrücken lässt.

$$T(\omega) \vec{u}(\omega) = \vec{f}(\omega) = \kappa(\omega) \vec{f}(\omega) \quad (4.63)$$

$$T(\omega) \vec{u}(\omega) = \vec{f}(\omega) \quad (4.64)$$

Nun ist man daran interessiert, wie $\kappa(\omega)$ die Wellenfelder beeinflusst. Hierzu wird Gleichung 4.64 mit $\kappa(\omega)$ multipliziert.

$$T(\omega) \vec{u}(\omega) \kappa(\omega) = \vec{f}(\omega) \kappa(\omega) = \vec{f}(\omega) \quad (4.65)$$

Ein Vergleich mit 4.63 und Multiplikation mit $T(\omega)^{-1}$ zeigt:

$$\vec{u}(\omega) \kappa(\omega) = \vec{u}(\omega) \quad (4.66)$$

Die Wellenfelder ändern sich durch die Manipulation $\kappa(\omega)$ in gleicher Weise wie das Signal. Die Suche nach dem optimalen Quellsignal lässt sich nun als Minimierungsproblem beschreiben.

$$\min_{\kappa} \left(\vec{d}(\omega) - \vec{u}(\omega) \right) \quad (4.67)$$

Gesucht ist $\kappa(\omega)$, sodass die Abweichung zwischen dem observierten Wellenfeld $\vec{d}(\omega)$ und dem synthetisch modellierten Wellenfeld $\vec{u}(\omega)$ minimal ist. Dieses optimale Wellenfeld $\vec{u}(\omega)$ folgt aus dem optimalen Quellsignal $\vec{f}(\omega)$. Zur Verfügung steht aber die synthetische Beobachtung $\vec{u}(\omega)$ aus $\vec{f}(\omega)$, die mittels Minimierungsparameter $\kappa(\omega)$ und Gleichung 4.66 an das optimale Wellenfeld $\vec{u}(\omega)$ angepasst werden soll. Ist $\kappa(\omega)$ gefunden, so kann mit Gleichung 4.62 aus dem initial verwendeten Quellsignal $\vec{f}(\omega)$ das gesuchte optimale Signal $\vec{f}(\omega)$ berechnet werden.

Als Misfit kann prinzipiell jedes Maß dienen, das den Abstand zwischen observierten und synthetischen Daten quantifiziert. Die Verwendung der L2-Norm führt auf die aus Kapitel 4.3.2 bekannte Methode der kleinsten Fehlerquadrate. Aufgrund der linearen Abhängigkeit von $\kappa(\omega)$ lässt sich so die Inversion in einem Schritt, ohne Iterationen, durchführen.

Die L2-Norm der komplexen Datenabweichung lässt sich mit den komplex konjugierten Größen, gekennzeichnet durch *, ausdrücken.

$$S = \frac{1}{2} \sum_{i=1}^N [d_i(\omega) - \bar{u}_i(\omega)\kappa(\omega)] [d_i^*(\omega) - \bar{u}_i^*(\omega)\kappa^*(\omega)] \quad (4.68)$$

Das Minimum findet sich über die Nullstelle der Ableitung $\frac{\partial S}{\partial \kappa}$. Zur Handhabung des komplex konjugierten $\kappa^*(\omega)$ in der Ableitung erfolgt eine Aufspaltung in Real- und Imaginärteil ($\kappa = \kappa^R + i\kappa^I$).

$$\frac{\partial S}{\partial \kappa} = \frac{\partial S}{\partial \kappa^R} \frac{\partial \kappa^R}{\partial \kappa} + \frac{\partial S}{\partial \kappa^I} \frac{\partial \kappa^I}{\partial \kappa} = \frac{\partial S}{\partial \kappa^R} - i \frac{\partial S}{\partial \kappa^I} \quad (4.69)$$

Ausmultiplizieren von Gleichung 4.68 und Trennung des Real- und Imaginärteils von κ ergibt:

$$S = \frac{1}{2} \sum_{i=1}^N d_i d_i^* - d_i \bar{u}_i^* (\kappa^R - i\kappa^I) - d_i^* \bar{u}_i (\kappa^R + i\kappa^I) + \bar{u}_i \bar{u}_i^* (\kappa^R \kappa^R + \kappa^I \kappa^I) \quad (4.70)$$

Nach Ausführung der Ableitung mit 4.69 und Nullsetzen

$$\frac{\partial S}{\partial \kappa} = \frac{1}{2} \sum_{i=1}^N (-d_i \bar{u}_i^* - d_i^* \bar{u}_i + 2\bar{u}_i \bar{u}_i^* \kappa^R) - i (i d_i \bar{u}_i^* - i d_i^* \bar{u}_i + 2\bar{u}_i \bar{u}_i^* \kappa^I) \quad (4.71)$$

$$= -d_i^* \bar{u}_i + \bar{u}_i \bar{u}_i^* \kappa^R - i \bar{u}_i \bar{u}_i^* \kappa^I = -d_i^* \bar{u}_i + \bar{u}_i \bar{u}_i^* \kappa^* = 0 \quad (4.72)$$

folgt schließlich das gesuchte κ . Mit Gleichung 4.62 kann so das optimale Signal $\vec{f}(\omega)$ berechnet werden.

$$\kappa = (\bar{u}_i \bar{u}_i^*)^{-1} \bar{u}_i^* d_i \quad (4.73)$$

Dieses Ergebnis unterscheidet sich zu der in Pratt (1999) angegebenen um die komplexe Konjugation.

Liegen die Wellenfelder \vec{d} , \vec{u} und das initial verwendete Quellsignal \vec{f} im Zeitbereich vor, so können sie in den Frequenzraum transformiert werden, um κ mit Gleichung 4.73 und das gesuchte Signal mit Gleichung 4.62 zu bestimmen. Rücktransformation ergibt endgültig $\vec{f}(t)$. Dies bedeutet auch, dass eine Vorwärtsmodellierung ausreichend für die Inversion des Quellsignals ist.

4.8. Inversionsschema

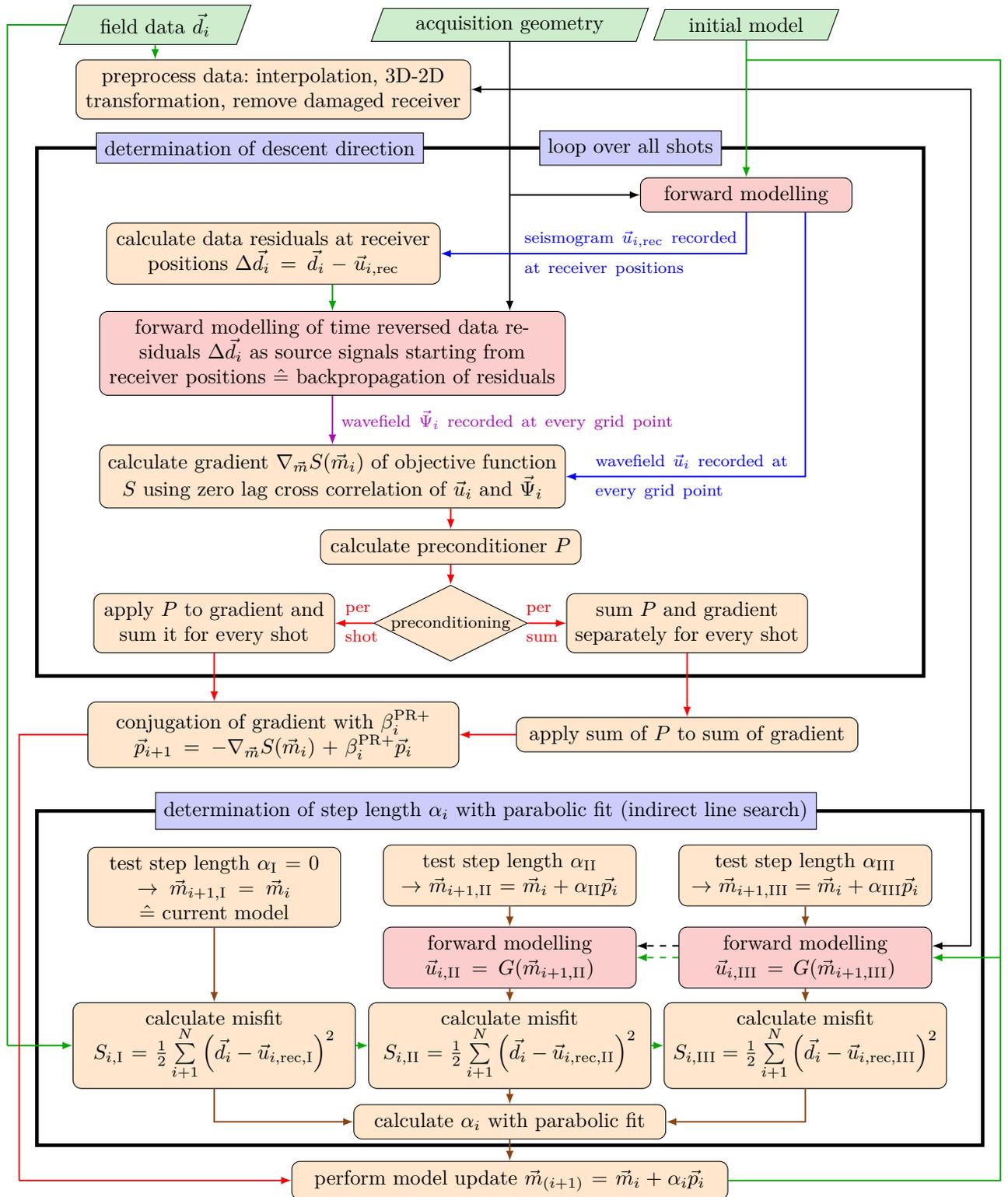


Abbildung 4.8.: Ablaufdiagramm der Wellenforminversion.

5. Beschreibung der Felddaten

Zur Durchführung der Wellenforminversion werden reale Daten verwendet, welche in anonymisierter Form für wissenschaftliche Zwecke zur Verfügung gestellt werden. Zweck der Messkampagne war die 3-dimensionale Charakterisierung eines marinen Kohlenwasserstoffreservoirs, das hier jedoch nicht von Interesse ist.

5.1. Akquisitionsgeometrie

Insgesamt wurden 20 sogenannte *ocean bottom cables* (OBC) mit einer Länge von je 6000 m parallel zueinander am Meeresgrund verlegt. Sie sind in regelmäßigen Abständen mit 4-Komponenten Sensoren (drei Geophone und ein Hydrophone) besetzt. Entlang jeder der 20 Empfängerketten finden sich 61 Schusspunkte 6 m unterhalb der Wasseroberfläche. Hier wurden 36 Druckluftkanonen, angeordnet in einem ebenen 2D-Array von 15 m auf 16 m, zeitgleich gezündet. Da die Ausdehnung der Quellpunkte klein gegenüber der betrachteten Wellenlängen ist, wird jeder Schuss als einzelner Punkt idealisiert. Eine ausführliche Beschreibung zur seismischen Datenakquisition findet sich in [Drijkoningen \(2003\)](#).

Tabelle 5.1.: Übersicht zu den Akquisitionsparameter der Felddaten

Parameter	Wert
Aufzeichnungsdauer	6 s
Abtastrate	2 ms
Anzahl der Hydrophone	239
Länge des OBC	6000 m
Hydrophonabstand	25 m
Abstand der Schusspunkte	200 m
Tiefe der Schusspunkte	6 m
Anzahl der Quellen	61
Synchronisationstoleranz der Einzelquellen	1.5 ms
Luftdruck	138 bar $\pm 10\%$

Als Rohdaten liegen lediglich die von den Hydrophonen aufgezeichneten zeitlichen Differentiale des Drucks einer Empfängerkette mit 61 Schusssektionen vor. Aus technischen Gründen sind die Hydrophone und Schusspunkte nicht exakt entlang einer Linie ausgerichtet. Ihre relativen Positionen sind in Abbildung 5.1 dargestellt. Die Abweichungen von einer gemeinsamen Ebene sind verglichen mit den Wellenlängen gering. Daher wird für die gesamte Inversion die Projektion der Positionen in die y - z Ebene verwendet, was eine rein 2-dimensionale Betrachtung ermöglicht aber auch eine 3D-2D Transformation der Felddaten notwendig macht (s. Kapitel 6.2).

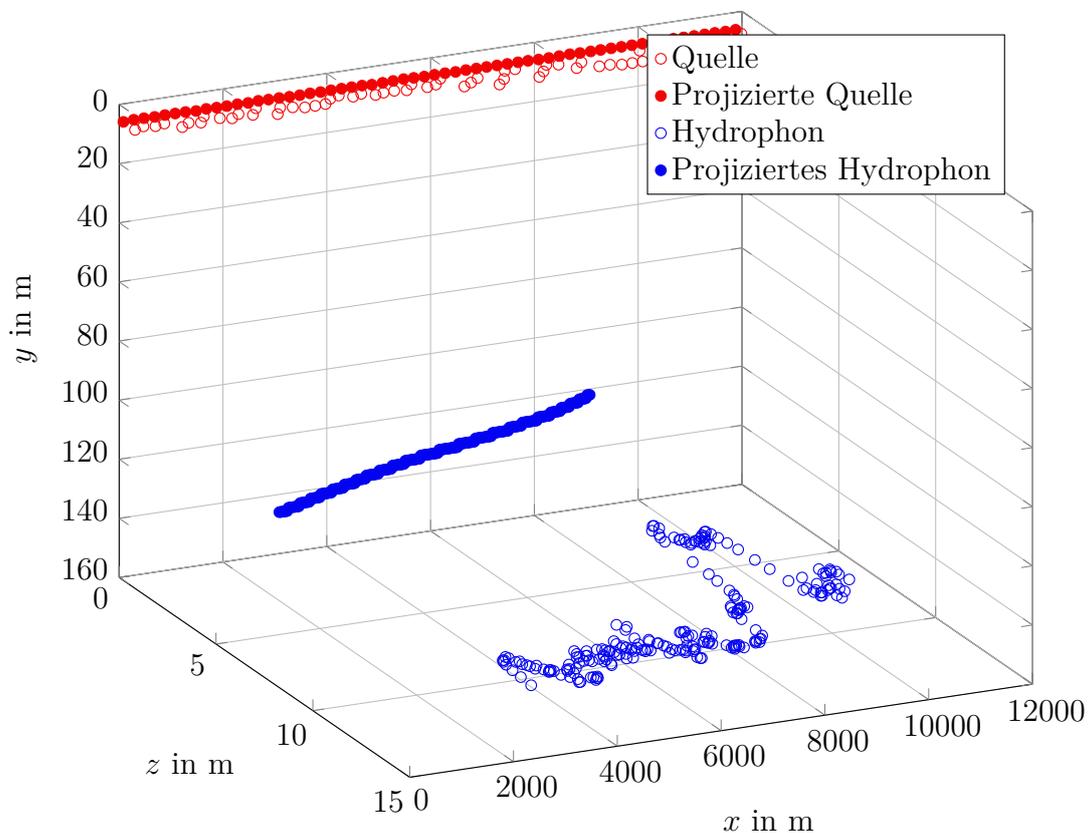


Abbildung 5.1.: Akquisitionsgeometrie der verwendeten Felddaten (relativ). In der Inversion wird die Projektion in die y - z Ebene verwendet. Für jede der 61 Quellen liegt eine Seismogrammsektion mit 239 Spuren vor. Die Nummerierung der Quellen erfolgt mit ansteigenden x -Werten.

5.2. Frequenzgehalt

Ein wesentlicher Aspekt der Felddaten ist deren Frequenzgehalt. Hoch- und niederfrequente Anteile können in der Inversion Verwendung finden. Hohe Frequenzen erfahren bezüglich dem Ausbreitungsweg eine stärkere Dämpfung als niedrige Frequenzen (vgl. Kapitel 2.4). Folglich sinkt das Signal-Rausch Verhältnis mit zunehmender Frequenz. Andererseits steigt mit der Frequenz die Empfindlichkeit der Wellen auf kleinräumige Kontraste im Medium. Das Auflösungsvermögen beträgt ca. $\frac{\lambda}{2}$, wobei eine wesentlicher Einfluss durch die Akquisitionsgeometrie besteht. Weiterhin wird die Inversion bei der *multiscale* Methode (s. Kapitel 7.3) mit Frequenzbänder durchgeführt, deren Frequenzen mit den Iterationsschritten zunehmen.

Ein weiterer Grund für eine obere Begrenzung der Frequenzen liegt in der Vorwärtsrechnung mit Finiten Differenzen, die praktisch den gesamten Rechenaufwand darstellt. Mit den minimalen simulierten Wellenlängen λ_{min} sinkt der notwendige Gitterabstand, um numerische Gitterdispersion zu unterdrücken (vgl. Gl. 3.6), was den Rechenaufwand t_R auf einem zweidimensionalen Gitter im Verhältnis $t_R \sim \lambda_{min}^2$ erhöht.

Der normierte Frequenzgehalt von Schuss 8 ist in Abbildung 5.2 dargestellt. Abbildung 5.3 zeigt beispielhaft die Seismogrammsectionen von Schuss 8, 19 und 31.

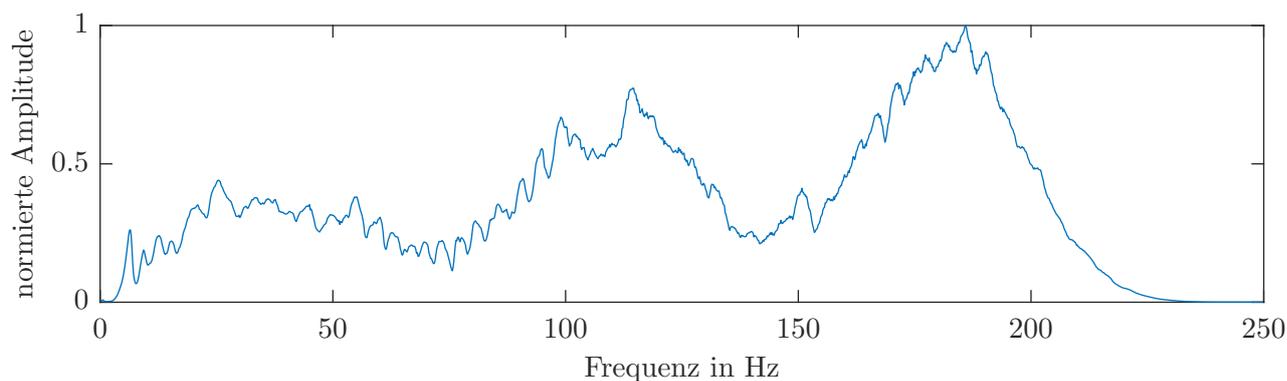


Abbildung 5.2.: Normiertes Frequenzspektrum der Seismogrammsection des 8. Schusses (vgl. Abbildung 5.3).

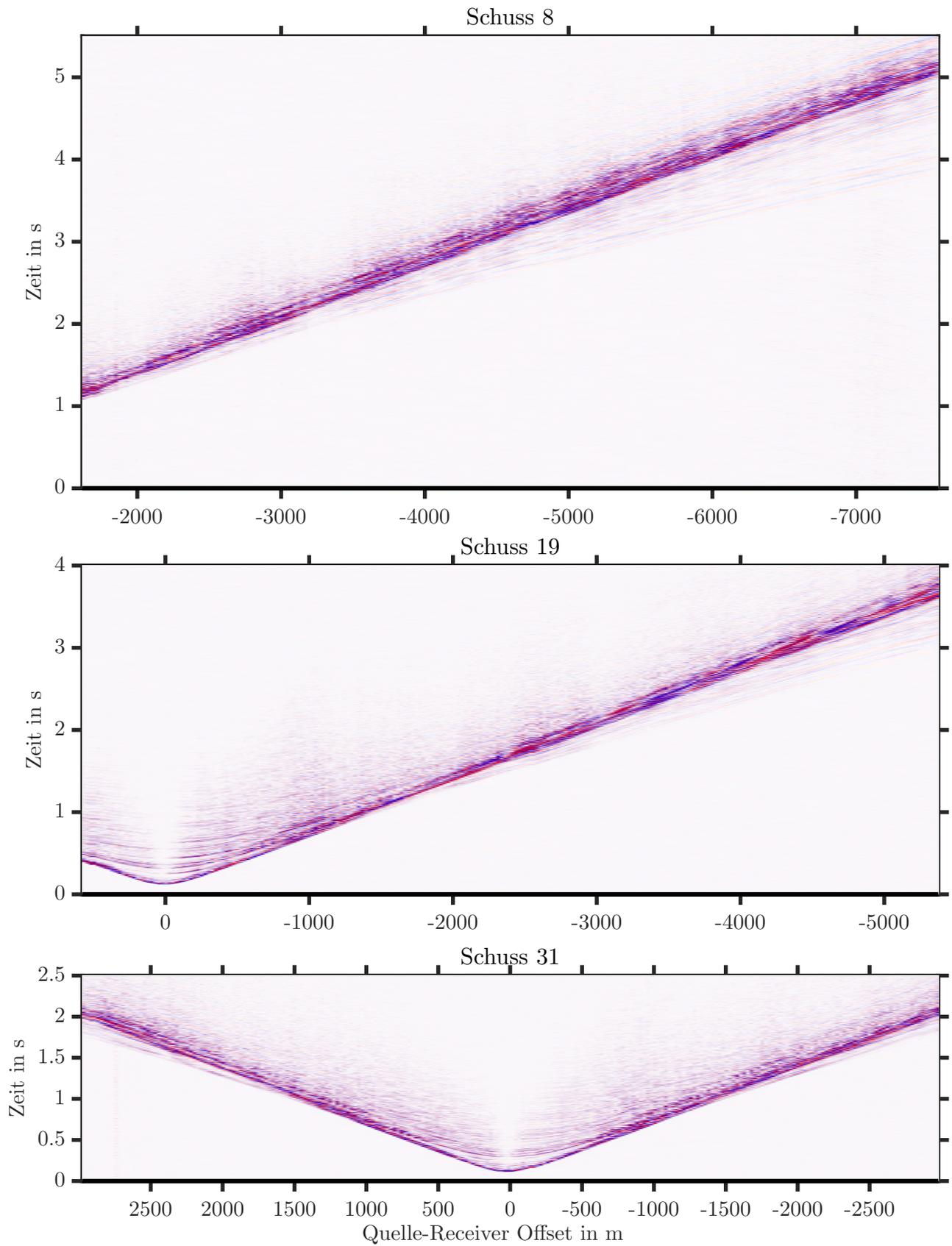


Abbildung 5.3.: Spurnormierte Seismogrammsektionen von Schuss 8, 19 und 31.

5.3. Geführte Wellen

Das Auftreten von geführten Wellen ist typisch für seismische Messungen in flachen Gewässern (Klein et al., 2005). Es ist daher interessant deren Existenz in einem slant-stack (Langsamkeits-Frequenz-Spektrum) anhand der Felddaten zu zeigen und später mit den synthetischen Seismogrammen der Inversion zu vergleichen. Aus Gleichung 2.23 geht hervor, dass die geführten Wellen sensitiv auf die Scherwellengeschwindigkeit des Meeresbodens reagieren. In dieser Arbeit wird eine rein akustische Inversion durchgeführt (akustische Vorwärtsmodellierung und Inversion nach v_P -Modell), d.h. die Ausbreitung von Scherwellen, sowie jegliche Konversionen werden ignoriert. Aus diesem Grund ist nicht zu erwarten, dass die Inversionsergebnisse die geführten Wellen in den Felddaten rekonstruieren können.

Der slant-stack von Schussektion 8 ist ein Abbildung 5.4 einzusehen. Es sind mehrere Moden deutlich zu erkennen. Ein weiteres Merkmal der geführten Wellen ist die Konvergenz der Dispersionsäste für hohe Frequenzen gegen die Kompressionswellengeschwindigkeit der Wasserschicht, was aus Gleichung 2.23 abgeleitet werden kann. Die durch das Startmodell und dem Taper auf $1520 \frac{\text{m}}{\text{s}}$ festgelegte Geschwindigkeit entspricht einer Phasenlangsamkeit von $6.58 \times 10^{-4} \frac{\text{s}}{\text{m}}$ und damit gut dem abzulesenden Wert.

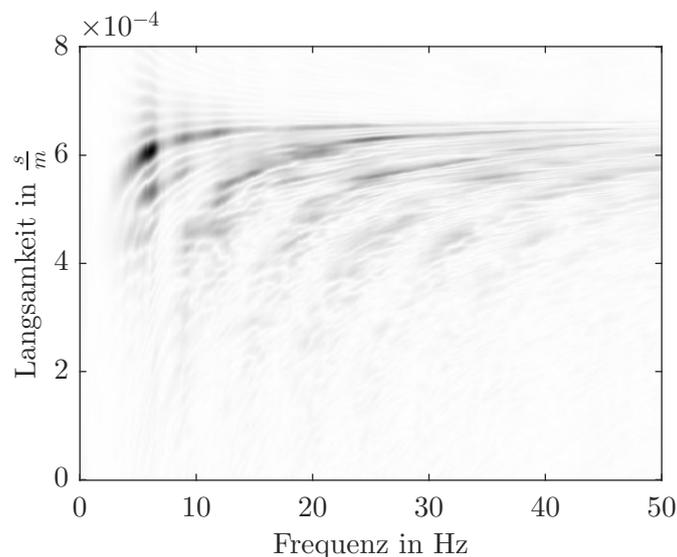


Abbildung 5.4.: Langsamkeits-Frequenz-Spektrum anhand der präprozessierten Felddaten von Schussektion 8.

6. Präprozessierung

Bevor die Felddaten für die Inversion verwendet werden können, müssen bestimmte Manipulationen vorgenommen werden. Dazu zählen die 3D-2D Transformation und die Interpolation. Daneben existieren eine Vielzahl an optionalen Eingriffen, meist mit dem Ziel die Datenqualität aufzubereiten. In diesen Bereich fallen Reduzierung des Rauschens, Frequenzfilterung, und das Ausschneiden von Seismogrammteilen, auf was in dieser Arbeit komplett verzichtet werden soll.

6.1. Qualitätskontrolle

Durch die Sichtung der Daten fiel im Zeit- sowie im Frequenzbereich auf, dass die Seismogramme mit Spurnummer 11 in jeder Schusssektion ein deutlich erhöhtes Rauschniveau aufweisen. Dies ist vermutlich auf ein beschädigtes Hydrophon zurückzuführen. Aus diesem Grund wird die Spur 11 aus den Daten entfernt, womit sich die Anzahl der effektiven Hydrophone von 240 auf 239 reduziert.

6.2. 3D-2D Transformation

Die Akquisition der Felddaten erfolgte durch ein Quellarray, das als idealisierte Punktquelle im 3-dimensionalen Raum betrachtet wird. Die Anpassung der synthetischen an die observierten Daten soll jedoch in zwei Dimensionen erfolgen, weshalb auch die Vorwärtsrechnungen in nur zwei räumlichen Dimensionen durchgeführt werden. Im Vergleich zu 3-D Wellenfeldsimulationen kann der Rechenaufwand so um mehrere Größenordnungen reduziert werden.

Der Vergleich zweier Wellenfelder einer Punktquelle, von dem eines in drei Dimensionen und das andere in zwei Dimensionen propagiert, zeigt, dass sie sich in Amplitude, aufgrund der unterschiedlichen geometrischen Dämpfung (vgl. Kapitel 2.4) und in Phase unterscheiden. Identische Wellenfelder erhält man durch Verwendung einer entlang

der dritten Dimension (senkrecht zur Betrachtungsebene) unendlich ausgedehnten Linienquelle. Um eine direkte 2-D Betrachtung von Felddaten zu ermöglichen, muss demnach eine Linienquelle in der Akquisition verwendet werden. Der hierfür erforderliche Mehraufwand ist allerdings nicht gerechtfertigt, denn mit sogenannten 3D-2D Transformationen ist es möglich Amplitude und Phase der Felddaten unter bestimmten Annahmen so zu manipulieren, dass sie einer Akquisition mit Linienquelle näherungsweise entsprechen.

Es existieren verschiedene Transformationen, die sich aus spezifischen Annahmen bezüglich des Modells ergeben. Es wird zwischen Einzelspurtransformationen, die unabhängig auf die Spuren einer Schusssektion angewendet werden aber nur für bestimmte Wellentypen (Reflexionen, Refraktionen, ...) gelten und Integraltransformationen, wo Seismogramme eines kompletten Profils benötigt werden, unterschieden. Weit verbreitet ist die Einzelspurtransformation nach [Pica et al. \(1990\)](#). Hier werden die Daten mit \sqrt{t} multipliziert (t : Laufzeit), was den geometrischen Ausbreitungsverlust auf das 2-D Maß korrigiert und eine Faltung (Konvolution) mit $\frac{1}{\sqrt{t}}$ (entspricht Phasenverschiebung um $\frac{\pi}{4}$) vorgenommen. Diese Transformation ist eine Näherung für homogene Medien und reflektierte Wellen.

In dieser Arbeit wird die für homogene Medien und reflektierte Wellen exakt gültige 3D-2D Transformation F_{3D-2D} verwendet, die als Faltung auf jede Spur angewendet wird ([Forbriger et al., 2014](#)). r ist der Quelle-Empfänger Abstand, t die Laufzeit, $v_W = 1500 \frac{\text{m}}{\text{s}}$ die Kompressionswellengeschwindigkeit in Wasser und Θ die Heaviside-Funktion.

$$F_{3D-2D} = \frac{2r}{\sqrt{t \left(t + \frac{2r}{v_W} \right)}} \Theta(t) \quad (6.1)$$

Die Auswirkungen der Transformation sind beispielhaft an drei Seismogrammen der Seismogrammsektion 8 in [Abbildung 6.1](#) gezeigt.

6.3. Interpolation der Felddaten

Die zeitliche Interpolation der Felddaten ist rein technischer Natur und hat keinen physikalischen Hintergrund. Dabei wird die ursprüngliche Abtastrate der Felddaten von 2 ms auf die der Simulation von 0.9 ms geändert. Dies ist notwendig, um beispielsweise den Misfit zwischen Felddaten und synthetischen Seismogrammen berechnen zu können.

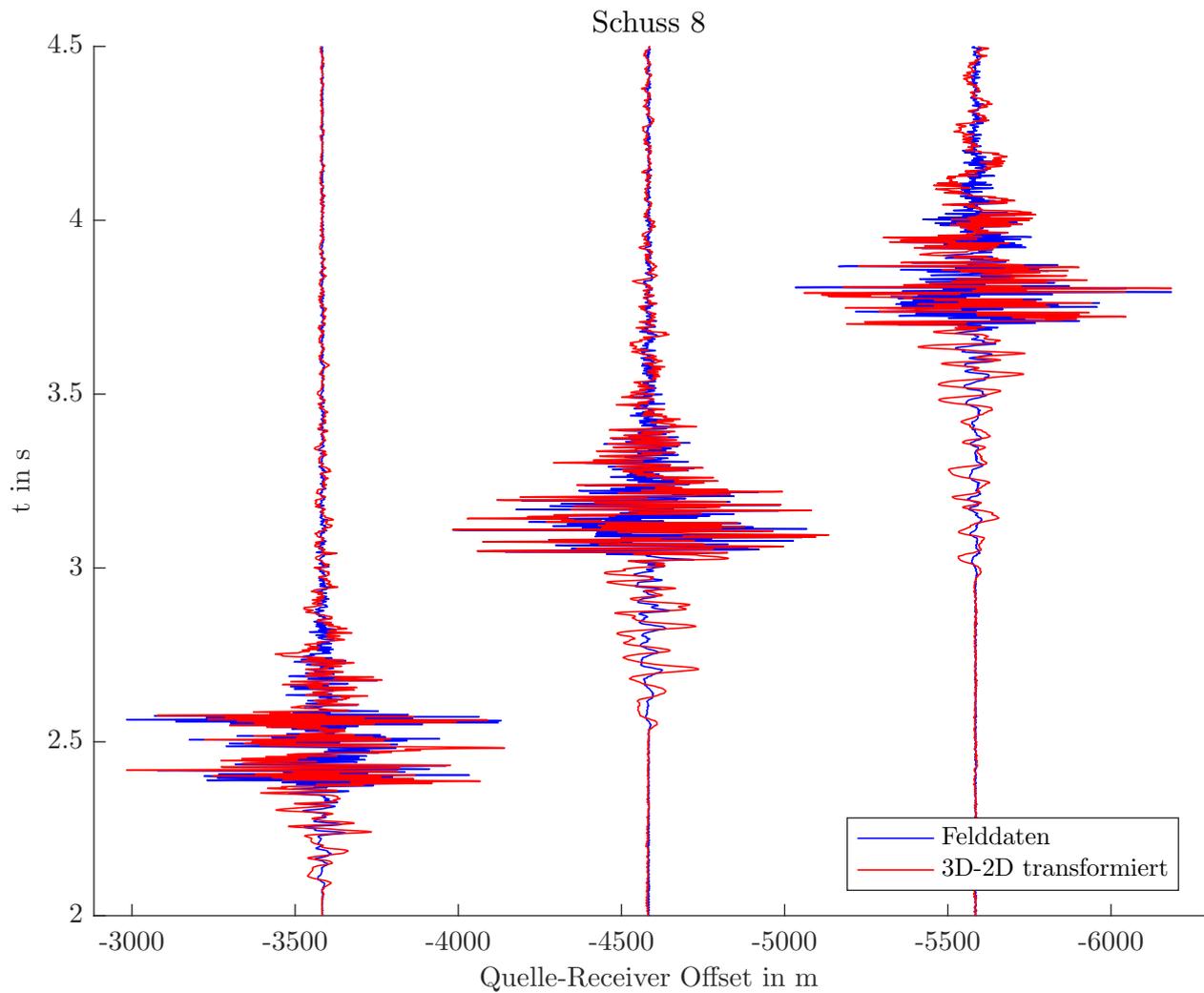


Abbildung 6.1.: Effekt der 3D-2D Transformation auf Seismogramme der Schusseksection 8 der Felddaten.

7. Wellenforminversion mit Referenzeinstellungen

In diesem Kapitel werden alle Parameter und Einstellungen einer Wellenforminversion aufgeführt, die als Referenz für nachfolgende Untersuchungen gilt (s. Tabelle 7.1). Zudem findet eine ausführliche Diskussion des Inversionverlaufs und der Ergebnisse statt. Im Anhang A.2 findet sich ein Video mit der Evolution der Inversionsergebnissen.

Die Intention der Referenzlösung besteht darin, mit möglichst geringem methodischen Aufwand, wie der Verwendung des konjugierten Gradienten als Abstiegsrichtung und minimalem Präprozessieren, ein möglichst gutes Inversionsergebnis zu erzeugen. Im folgenden Abschnitt wird das Startmodell vorgestellt.

Tabelle 7.1.: Referenzeinstellungen der Wellenforminversion

Vorwärtssimulation	
2-D finite Differenzen Lösung der akustischen Wellengleichung im Zeitbereich	
Genauigkeit (Zeit)	zweite Ordnung
Genauigkeit (Raum)	achte Ordnung
Gitterpunktabstand	5 m
Zeitschritt	0.9 ms
Inversion	
Zielfunktion	L2-Norm
Inversionsparameter	v_p -Modell
Präprozessierung	3D-2D Transformation, Interpolation
Abstiegsrichtung	präkonditionierter konjugierter Gradient
Präkonditionierung des Gradienten	inverse Diagonale der approx. Hessematrix, Taper
Anwendung der Präkonditionierung	per Summe des Gradienten
Schrittweitenbestimmung	parabolische Interpolation
Inversion des Quellsignals	pro Frequenzband
progressive Bandpassfilterung der Felddaten (<i>multiscale</i>)	
keine Glättung des Gradienten	

7.1. Startmodell

Als v_P -Startmodell dient das Ergebnis einer Laufzeit-tomographie des Untersuchungsgebiets, welches zusammen mit den Felddaten bereitgestellt wurde. Das Dichtemodell wird mithilfe der empirischen Relation nach Gardner (Gardner et al., 1974) aus dem v_P -Modell abgeleitet:

$$\rho = 310 v_P^{0.25} \quad \text{mit } [\rho] = \frac{\text{kg}}{\text{m}^3}, \quad [v_P] = \frac{\text{m}}{\text{s}} \quad (7.1)$$

Des Weiteren sind die bathymetrischen Messdaten des betrachteten 2D-Tiefenschnitts vorhanden. Für das Meerwasser wird eine Kompressionswellengeschwindigkeit von $v_P = 1520 \frac{\text{m}}{\text{s}}$ und eine Dichte von $\rho = 1000 \frac{\text{kg}}{\text{m}^3}$ angesetzt. v_P -Start- und Dichtemodell sind in Abbildung 7.1 dargestellt.

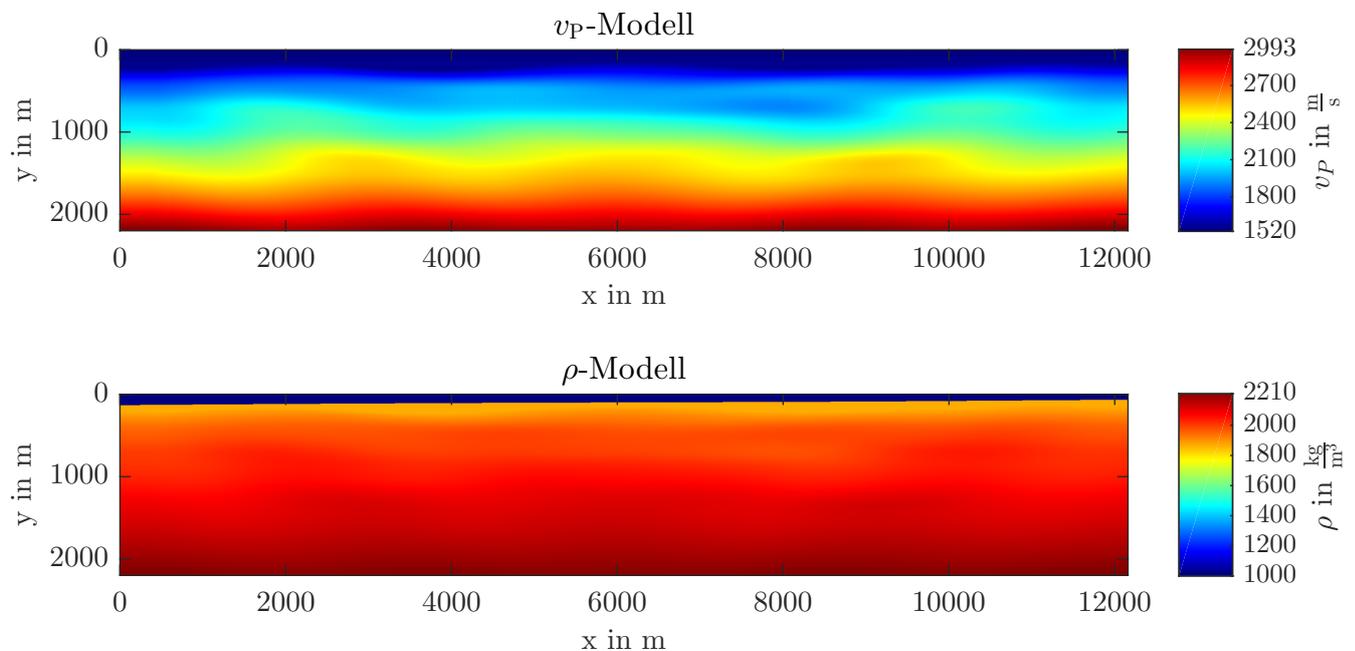


Abbildung 7.1.: v_P -Startmodell und das daraus abgeleitete Dichtemodell.

7.2. Taper

Das Wissen über die Meerestiefe kann man über ein Taper in die Inversion einfließen lassen. Der Taper ist ein Teil der Präkonditionierung des Gradienten. Dieser hat die Größe des Modells mit Werten zwischen 0 und 1 und wird mit dem finalen Gradienten jeder

Iteration multipliziert. Dadurch werden effektiv jegliche Modellupdates in der Wassersäule verhindert. Der verwendete Taper ist in Abbildung 7.2 einzusehen.

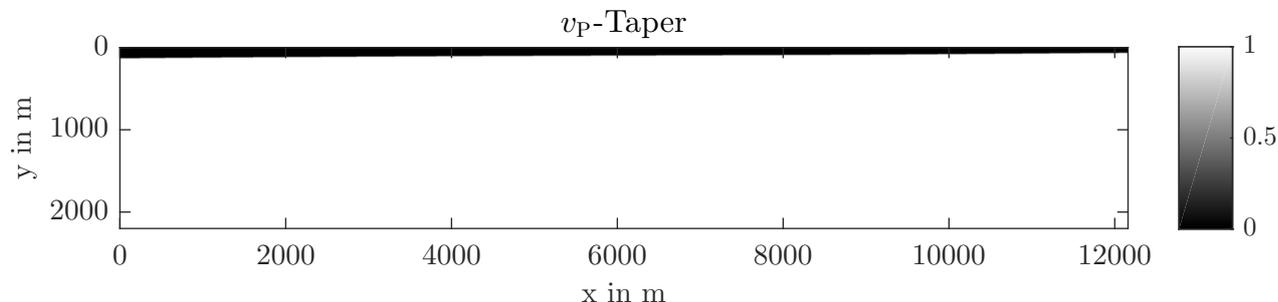


Abbildung 7.2.: v_P -Taper zur Unterdrückung von Modellupdates in der Wassersäule.

7.3. Mehrstufige Inversionsstrategie

Eine Möglichkeit die Nichtlinearität der Misfitfunktion $S(\vec{m})$ zu lindern, ist es für glattere $S(\vec{m})$ zu sorgen. Dies führt zu einem Inversionsalgorithmus, der weniger anfällig dafür ist von Nebenminima gefangen zu werden. Erreicht werden kann dies mit einer mehrstufigen Inversionsstrategie (*multiscale strategy*) (Bunks et al., 1995). Hierbei wird der Frequenzgehalt des Quellsignals mit fortschreitenden Iterationen mithilfe eines Bandpasses sukzessive erhöht. Zunächst werden damit lediglich die langwelligen Anteile propagiert und mit den ebenfalls gefilterten observierten Seismogrammen verglichen. Der Sprung in die nächste Frequenzstufe erfolgt, wenn die Misfitänderung relativ zur vorhergehenden Iteration einen definierten Wert unterschreitet und der betrachtete Frequenzgehalt somit erschöpfend behandelt wurde. Tabelle 7.2 führt die hier verwendeten 22 Stufen mit den beiden Grenzfrequenzen f_{low} und f_{high} sowie dem jeweiligen Abbruchkriterium Prog_S auf.

7.4. Beurteilungskriterien des Inversionserfolgs

Eine schnelle Bewertung des Inversionsverlaufs gewährt die Entwicklung der Zielfunktion. Sie zeigt, ob der Algorithmus in der Lage ist, aus Sicht der observierten Daten, bessere Modelle zu erzeugen.

Tabelle 7.2.: Inversionsstufen der *multiscale* Strategie mit der sich progressiv erhöhenden oberen Grenzfrequenz f_{high} , der unteren Grenzfrequenz f_{low} und dem Abbruchkriterium Prog_S .

Stufe	f_{low} Hz	f_{high} Hz	Prog_S	Stufe	f_{low} Hz	f_{high} Hz	Prog_S
1	2	4	0.01	12	2	15	0.01
2	2	5	0.01	13	2	16	0.01
3	2	6	0.01	14	2	17	0.01
4	2	7	0.01	15	2	18	0.01
5	2	8	0.01	16	2	19	0.01
6	2	9	0.01	17	2	20	0.01
7	2	10	0.01	18	2	21	0.01
8	2	11	0.01	19	2	22	0.01
9	2	12	0.01	20	2	23	0.01
10	2	13	0.01	21	2	23	0.001
11	2	14	0.01	22	2	23	0.0001

Da die Zielfunktion die Differenz zwischen observierten und synthetischen Daten als Skalar quantifiziert, ist es überaus hilfreich die exakte Herkunft der Abweichungen in den Seismogrammen zu ergründen. Hier wird sichtbar, ob es sich um Amplituden- oder Phasenfehler handelt.

In Kapitel 5.3 konnte die Existenz von geführten Wellen in den Felddaten gezeigt werden. Ein slant-stack auf Basis der synthetischen Seismogramme des finalen v_P -Modells wird zeigen inwieweit die Moden der geführten Wellen rekonstruiert werden können. Dies ist nicht selbstverständlich, da sie sensitiv auf die Scherwellengeschwindigkeit des Meeresbodens reagieren und eine rein akustische Inversion durchgeführt wird.

Ein weiteres Kriterium bietet ein Vergleich der invertierten Quellsignale jedes Schusses. Da stets das gleiche Quellsignal verwendet wurde, sollten sich die Signale weitestgehend decken.

Anders als bei Inversionsexperimenten mit synthetisch erzeugten 'Felddaten', ist das wahre v_P -Modell bei der Realdatenanwendung nicht bekannt, was die Bewertung des Inversionsergebnisses anhand des resultierenden Modells zunächst unmöglich macht. Neben den Felddaten selbst wurde auch ein Migrationsergebnis bereit gestellt. Dieses kann nicht als

wahres Modell angesehen werden, was aber auch nicht erforderlich ist. Übereinstimmende Merkmale weisen stark auf tatsächliche Inhomogenitäten hin, wenn solch grundsätzlich verschiedene Methoden zu gleichen Resultaten führen.

7.5. Bewertung der Ergebnisse

Die Evolution des Misfits in Abbildung 7.3 zeigt einen für eine erfolgreiche Wellenforminversion typischen Verlauf. Bis hin zu einer oberen Grenzfrequenz von 6 Hz (Iteration 48) fällt der Misfit kontinuierlich auf ein Minimum von ca. 0.1. Die zugehörigen Seismogramme in Abbildung 7.4, beispielhaft für den 8. Schuss, machen deutlich, dass die Datenanpassung sowohl in Phase als auch in Amplitude gelingt.

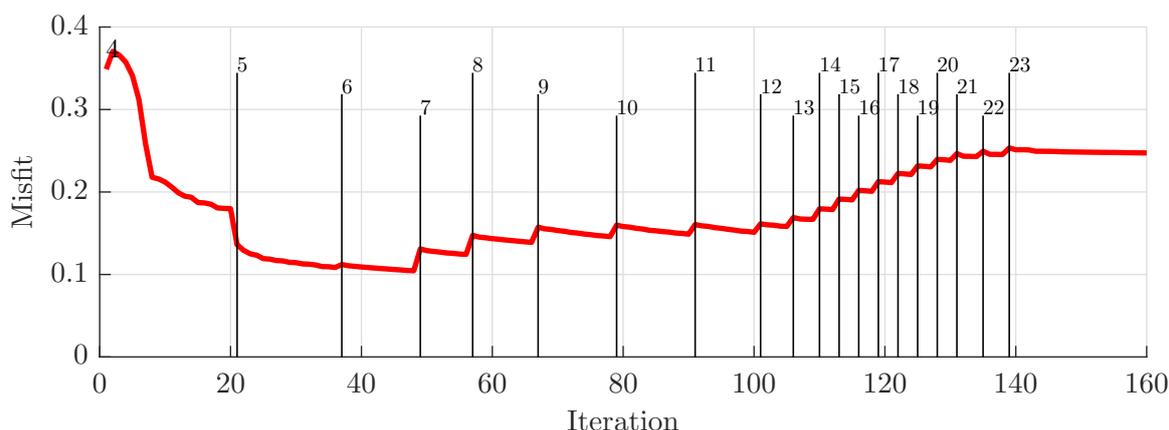


Abbildung 7.3.: Entwicklung des Misfit der Referenzlösung über dem Iterationsschritt (rot). Die senkrechten Linien kennzeichnen die Erhöhung der oberen Grenzfrequenz (4 Hz - 23 Hz) des Bandpassfilters, der vor jeder Iteration auf die observierten Daten angewendet wird.

Mit fortschreitenden Iterationen kommt es bei jeder Erhöhung des Frequenzgehalts zu einem Sprung im Misfit. Das ist ein normales Verhalten, denn mit jeder Frequenzstufe steigt der Informationsgehalt, der dem Algorithmus zur Verfügung gestellt wird. Dieser benötigt einige Iterationen das Model entsprechend anzupassen. Trotzdem ist zu beobachten, wie der Misfit wieder deutlich ansteigt, was sich mit dem typischerweise schlechter werdenden Signal-Rausch-Verhältnis erklären lässt. Das heißt der Rauschanteil am Signal der Felddaten nimmt mit der Frequenz zu. Rauschen kann von der Inversion weder erkannt noch sinnvoll verarbeitet werden. Sehr wohl aber beeinflusst es das Inversionsergebnis und führt

zu nicht anpassbaren Datenabweichungen. So ist bei optischem Vergleich der Seismogramme in Abbildung 7.4 von Iteration 90 zu Iteration 160 eine deutliche Verschlechterung zu erkennen. Dass die Datenanpassung bei hohen Frequenzen schlechter ist als bei niedrigen, bedeutet noch keine Verschlechterung des resultierenden v_P -Modells. Hier steht der teilweise fehlgeschlagene Versuch der Datenanpassung der Tatsache gegenüber, die höheren Frequenzen bei der Misfitberechnung gänzlich unbeachtet zu lassen.

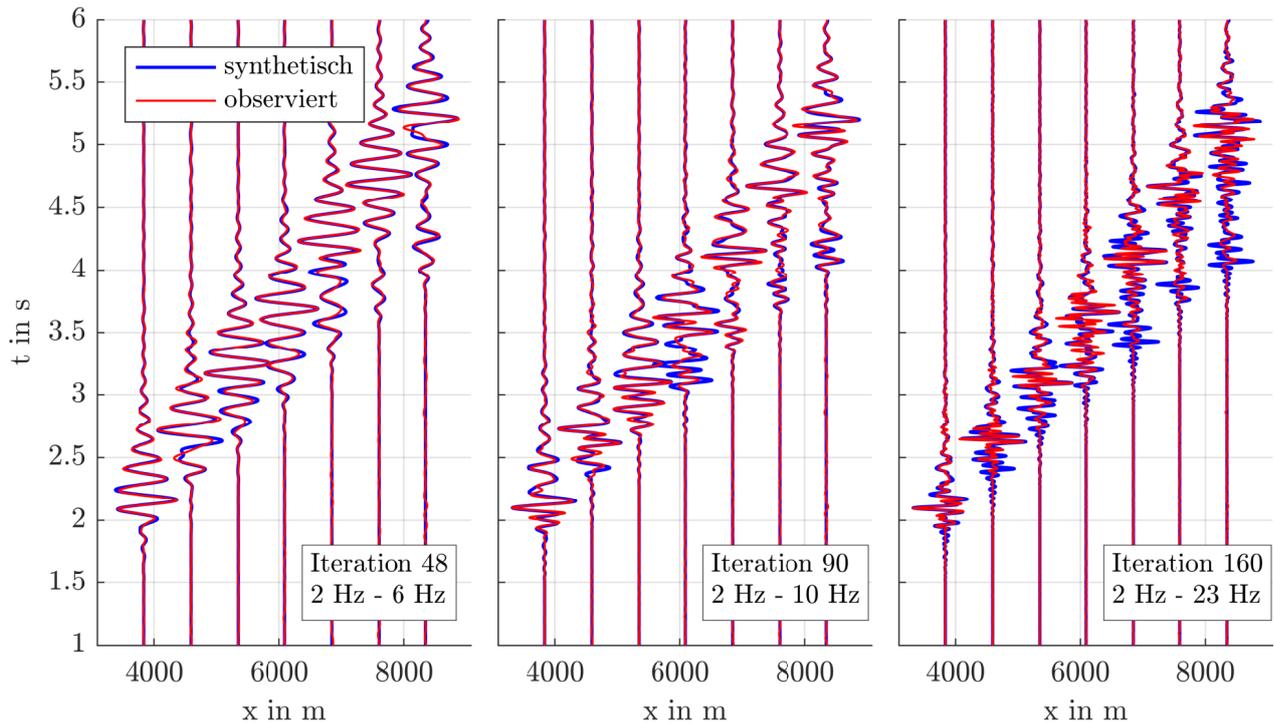


Abbildung 7.4.: Vergleich zwischen beobachteten und synthetischen Seismogrammen des 8. Schusses für verschiedene Iterationen der Referenzlösung. Der zugehörige Misfit und die Frequenzstufen können in Abbildung 7.3 abgelesen werden.

Das Langsamkeits-Frequenz-Spektrum der synthetischen Daten des finalen Modells (160. Iteration) von Schusssektion 8 ist links in Abbildung 7.5 zu sehen. Zu beachten ist, dass im Zusammenhang mit der *multiscale* Methode der Frequenzgehalt während der Inversion durch einen Bandpass begrenzt wird. Bei Iteration 160 wurde die maximale obere Grenzfrequenz von 23 Hz erreicht. Dies ist auch der Grund für das im Vergleich zu den Felddaten in Abbildung 5.4 abrupte Abreisen des Spektrums. Eine detaillierte Gegenüberstellung ermöglicht der rechte Teil von Abbildung 7.5. Hier sind alle Anteile, die nur in den Felddaten vorhanden sind grün und die Anteile, die nur in den synthetischen Daten vorhanden sind violett dargestellt. Bereiche mit ähnlicher Intensität sind als Graustufen

gekennzeichnet. Wie bereits erwähnt, ist nicht zu erwarten, dass das akustische Modell die geführten Wellen rekonstruiert, da sie sensitiv auf die Scherwellengeschwindigkeit des Meeresbodens reagieren. Entgegen dieser theoretischen Annahme können die Moden bis zur Schranke von 23 Hz größtenteils von den Inversionsergebnissen rekonstruiert werden.

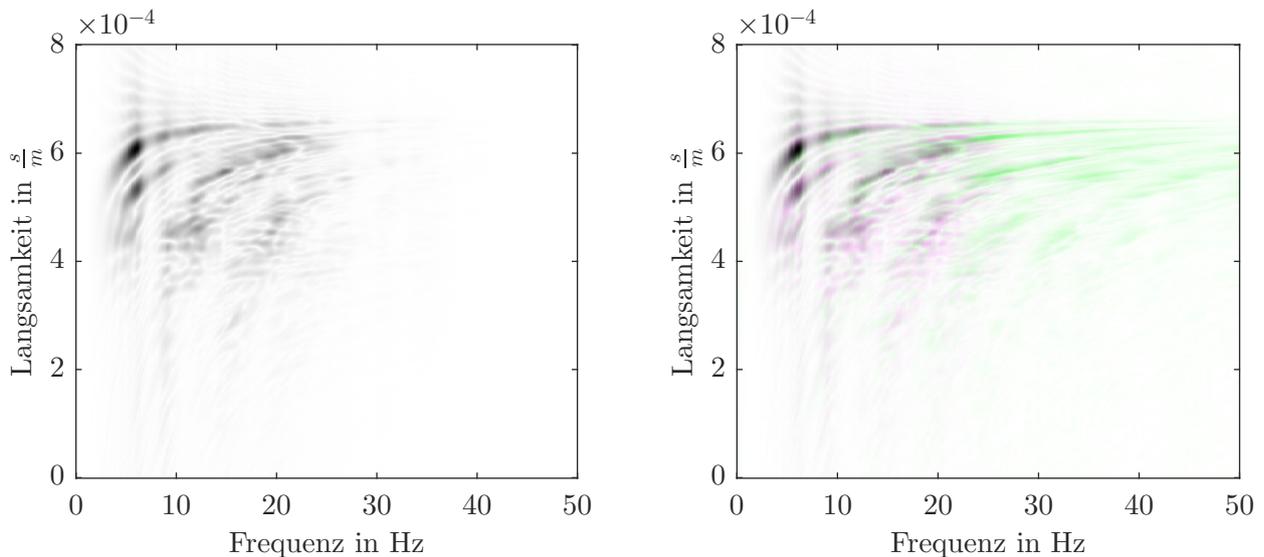


Abbildung 7.5.: Langsamkeits-Frequenz-Spektrum anhand der synthetischen Seismogramme des finalen Modells (160. Iteration) der Referenzlösung von Schussektion 8 (links) und Komposition dieses Spektrums mit dem der präprozessierten Felddaten (vgl. Abbildung 5.4) (rechts). Das Spektrum der Felddaten ist grün dargestellt, das der synthetischen Daten violett. Bereiche mit ähnlicher Intensität sind als Graustufen gekennzeichnet.

Die Inversion der Quellsignale wird zu Beginn jeder Frequenzerhöhung vorgenommen. Mit der Annahme einer hohen Wiederholgenauigkeit der Schussausführung sollten sich die invertierten Signale weitgehend decken. Hier muss jedoch beachtet werden, dass die Quellinversion den gesamten Unterschied zwischen observierten und synthetischen Daten durch ein optimiertes Signal versucht zu minimieren. Da ein Teil der Datenresiduen ihre Ursache stets in einem falschen Modell haben, ist das invertierte Quellsignal fehlerbehaftet. Dennoch kann Abbildung 7.6 weitgehend übereinstimmende Signalformen entnommen werden.

In Abbildung 7.7 sind verschiedene Modelle zur Beurteilung der Inversionsergebnisse aufgeführt. Ein Vergleich zwischen Startmodell a) und Inversionsergebnis b) zeigt, dass die langwelligen Strukturen durch die Inversion so beibehalten wurden, was auf ein geeignetes

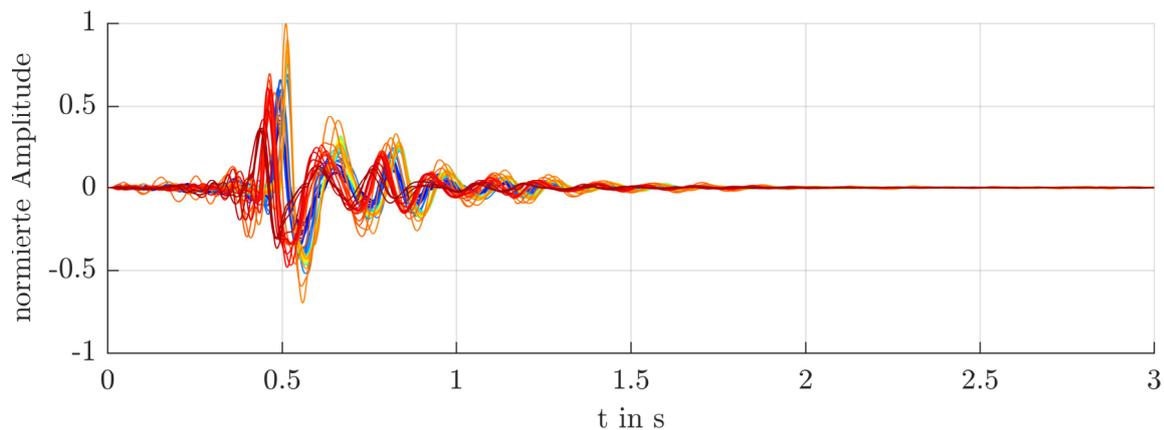


Abbildung 7.6.: Überlagerung der invertierten Quellsignale aller 61 Schüsse von der Referenzlösung nach 160 Iterationen.

Startmodell hinweist. Im Migrationsmodell c) sind Reflexionshorizonte zu sehen, die auf einige Verwerfungen hindeuten. Ein Großteil dieser Merkmale werden sowohl durch die Inversion als auch durch die Migration abgebildet (s. f) und e)), was als großen Erfolg verzeichnet werden kann.

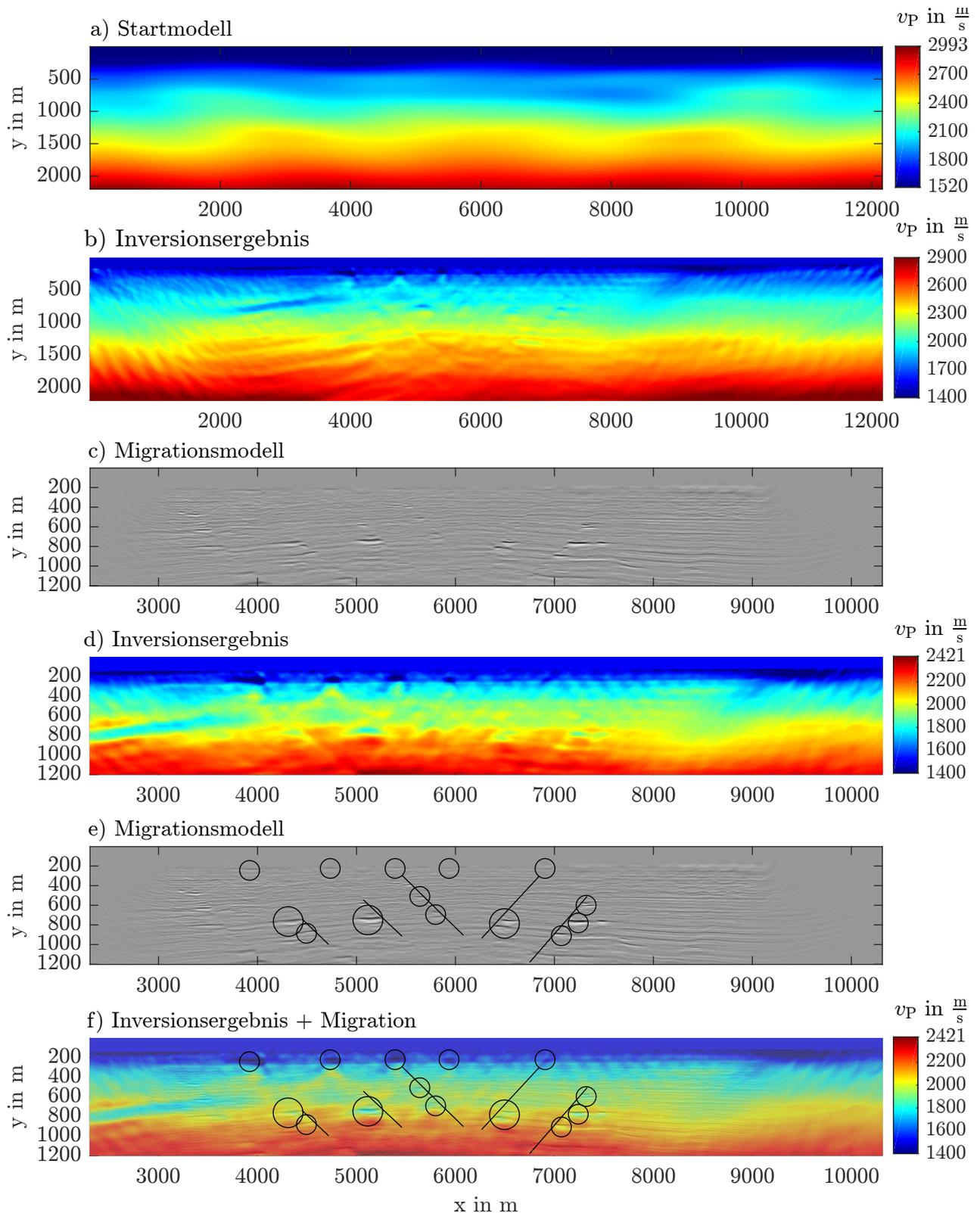


Abbildung 7.7.: Modelle zur Bewertung des Inversionsergebnisses der Referenzlösung. a) Startmodell und b) Inversionsergebnis nach 160 Iterationen der gesamten Ausdehnung. c) Migrationsmodell. d) Inversionsergebnis als Ausschnitt von b). e) und f) zeigen Merkmale, die sich sowohl im Inversions- als auch im Migrationsergebnis finden, wobei in f) beide überlagert sind.

8. Untersuchung von Quellsignalinversion und Präkonditionierung

Das zweite Ziel dieser Arbeit ist die Untersuchung von je zwei unterschiedlichen Vorgehen bei der Quellsignalinversion und der Präkonditionierung der Gradienten. Die Inversion des Quellsignals wird einerseits zu Beginn jedes neuen Frequenzbandes und andererseits jede Iteration durchgeführt. Die Präkonditionierung wird zwischen einer Anwendung per Summe aller Schüsse und per einzelner Schuss variiert. Generell wird der Präkonditionierer P für jeden Schuss berechnet. Bei der Anwendung per Summe werden P und Gradient, während der Schleife über alle Schüsse zur Berechnung der Abstiegsrichtung, separat aufsummiert. Anschließend wird die Summe von P auf die Summe der Gradienten angewendet. Für die alternative Anwendung per Schuss wird die Präkonditionierung direkt auf jeden einzelnen Schuss angewendet und das Ergebnis aufsummiert. Diese Unterscheidung kann, eingebettet in den gesamten Inversionsalgorithmus, im Ablaufdiagramm (Abbildung 4.8) eingesehen werden.

8.1. Motivation

Die in Kapitel 4.5.4 beschriebene Präkonditionierung nach [Shin et al. \(2001\)](#) soll zu einer Reduzierung der Quellartefakte führen. Trotz deren Verwendung zeigen sich in den Randbereichen des invertierten v_p -Modells deutliche Spuren der Quellen. Eine Vermutung ist, dass es durch die Präkonditionierung per Summe zu einer Verwässerung in der Wirkung der einzelnen Präkonditionierer P_i auf den zugehörigen Gradienten kommt. Dies veranschaulicht folgender schematischer Vergleich zwischen der Anwendung der Präkonditionierung per Summe und per Schuss (k_i : konjugierter Gradient).

$$(P_1 + P_2 + P_3 + \dots)(k_1 + k_2 + k_3 + \dots) \neq P_1k_1 + P_2k_2 + P_3k_3 + \dots \quad (8.1)$$

8.2. Beschreibung der Ergebnisse

Es wurde für jede der vier Kombinationen eine Inversion durchgeführt. Alle sonstigen Parameter und Einstellungen entsprechen denen der Referenzinversion aus Kapitel 7 und können in Tabelle 7.1 eingesehen werden. Die Referenzinversion findet sich unter diesen vier Varianten wieder. Der Fokus dieses Kapitels liegt auf Beschreibung und Interpretation der Unterschiede in den Inversionsergebnissen.

Die Evolution des Datenmisfits (Abbildung 8.1) zeigt bei allen vier Inversionen einen ähnlichen Verlauf, jedoch lassen sich auch einige Unterschiede ausmachen. So ist der Misfit bei der Kombination Präkonditionierung per Summe und Quellsignalinversion in jeder Iteration geringer als bei den übrigen Varianten. Des Weiteren wird das letzte Frequenzband von 23 Hz bei Präkonditionierung per Summe um 20 Iterationen später erreicht. Dies ist auf einen längeren Konvergenzvorgang in den Frequenzbändern bis ca. 12 Hz zurückzuführen. Das Abbruchkriterium für eine Erhöhung der Filterfrequenz ist über die relative Misfitänderung zwischen aktuellem und vorhergehendem Misfit definiert. Liegt sie unter 1 % wird ins nächste Frequenzband gewechselt. Eine längere Konvergenz ist nicht grundsätzlich negativ zu bewerten, da dies noch nichts über den Wert des Misfits aussagt, der angestrebt wird.

Eine Charakterisierung des Misfits anhand der Seismogramme gestaltet sich schwierig, da die Unterschiede sehr gering sind und sich keine Muster zeigen. Um einen fairen Vergleich der vier Inversionen hinsichtlich der Ergebnisqualität zu gewährleisten sind in Abbildung 8.2 die Iterationen der Seismogrammvergleiche so gewählt, damit jede Inversion die gleichen Frequenzbänder (bis 16 Hz) durchlaufen hat. In dieser 8. Schussektion sind einige Unterschiede in den Amplituden zu Beginn der Signale sichtbar. Doch auch hier unterscheiden sich die Datenabweichungen in Summe kaum voneinander, wie in Abbildung 8.1 abgelesen werden kann.

Das Interesse an einem Vergleich der invertierten Quellsignale ist in erster Linie durch die beiden unterschiedlichen Strategien bei der Signalinversion gegeben. Für die Durchführung der Quellsignalinversion in jeder Iteration können engere Signalverläufe als bei der einmaligen Durchführung pro Frequenzband erwartet werden, falls die realen Signale mit einer hohen Wiederholgenauigkeit ausgeführt wurden. Ein Vergleich der Signalverläufe in Abbildung 8.3 bestätigt diese Vermutung. Sowohl bei Präkonditionierung per Summe

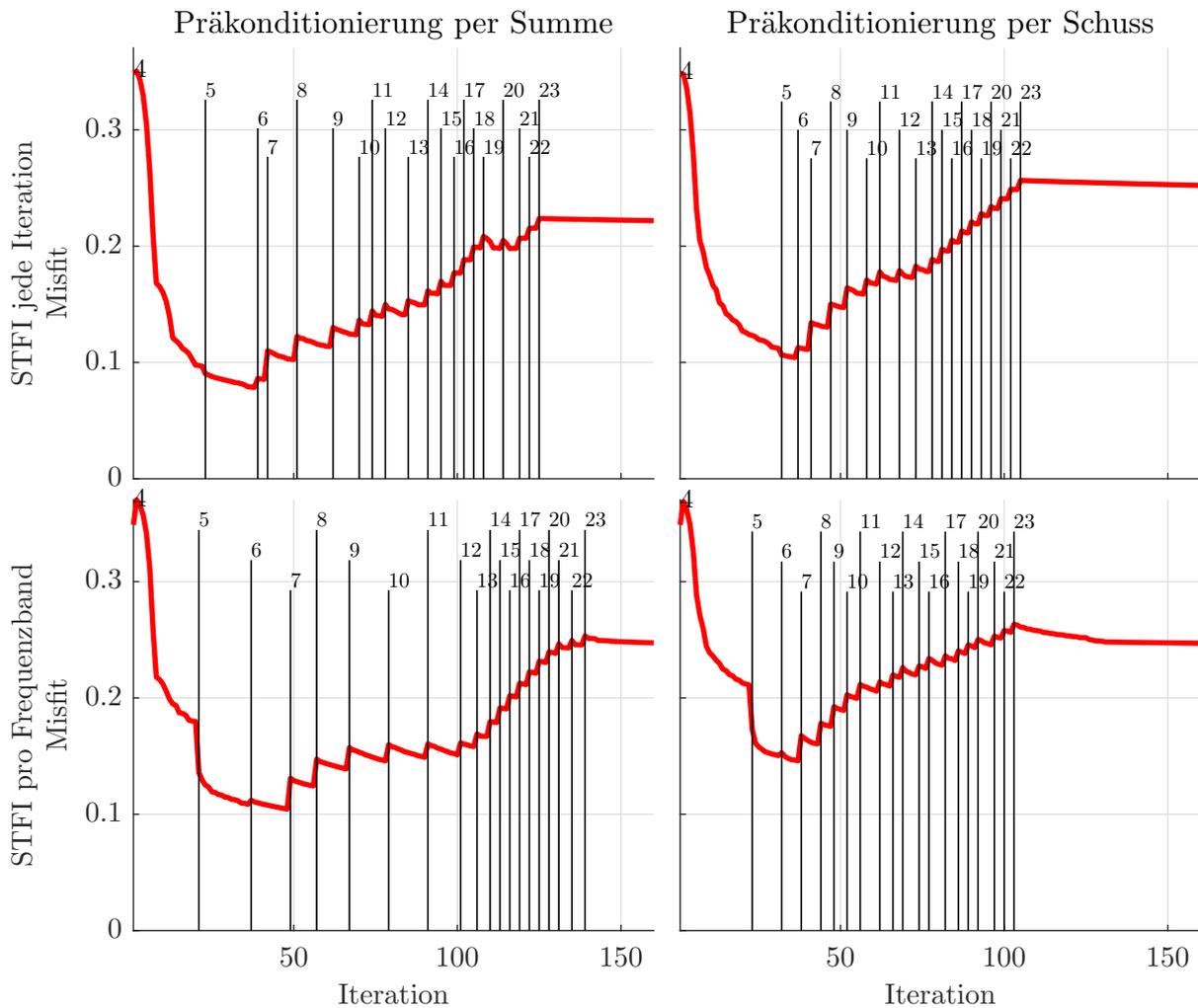


Abbildung 8.1.: Entwicklung des Misfits aller vier Inversionen über dem Iterationsschritt (rot). Die senkrechten Linien markieren die Iterationen bei denen die obere Grenzfrequenz (4 Hz - 23 Hz) des Bandpassfilters erhöht wird. Dieser Filter wird vor jeder Iteration auf die observierten Daten angewendet.

als auch bei Präkonditionierung per Schuss wirkt sich die häufigere Quellsignalinversion positiv aus. Besonders eng liegen die Signale bei der Kombination aus Präkonditionierung per Summe und Quellsignalinversion jede Iteration. Wie bereits im Theorieteil in Kapitel 4.7 angeführt, stehen die Inversion nach den Modellparametern und nach dem Quellsignal gewissermaßen in Konkurrenz zueinander. Beide versuchen die Abweichung zwischen observierten und synthetischen Daten zu minimieren. Einerseits durch Optimierung des Quellsignals und andererseits durch Optimierung der Modellparameter. Folgende Überlegung zeigt, dass die Qualität der beiden Inversionen voneinander abhängig sind. Je schlechter das synthetische Quellsignal an das reale angepasst ist, desto mehr Datenresiduen bleiben erhalten. Der nachfolgende Inversionsalgorithmus versucht alle Datenre-

siduen durch Anpassungen am v_P -Modell zu minimieren. Wobei die Anteile, welche auf das fehlerhafte Quellsignal zurückzuführen sind, zwangsläufig zu Fehler im v_P -Modell führen. Dieser Zusammenhang gilt auch für die umgekehrte Reihenfolge. Er kann durch die Beobachtungen in Abbildung 8.1 und 8.3 bestätigt werden.

Das zentrale Ergebnis der Untersuchung zeigt sich bei einem Vergleich der invertierten v_P -Modelle (Abbildung 8.4). Die Präkonditionierung per Schuss führt wie erwartet zu einer besseren Skalierung der Gradienten. Präziser ausgedrückt ist die von [Shin et al. \(2001\)](#) beschriebene Wirkung der Präkonditionierung deutlicher sichtbar. Der Gradient wird auf dem Ausbreitungsweg der Wellen von Quelle zu Receiver abgesenkt. Somit werden die Quellartefakte in Ihrer Ausprägung erheblich reduziert. Die beiden Varianten der Quellsignalinversion haben kaum Einfluss auf das v_P -Modell.

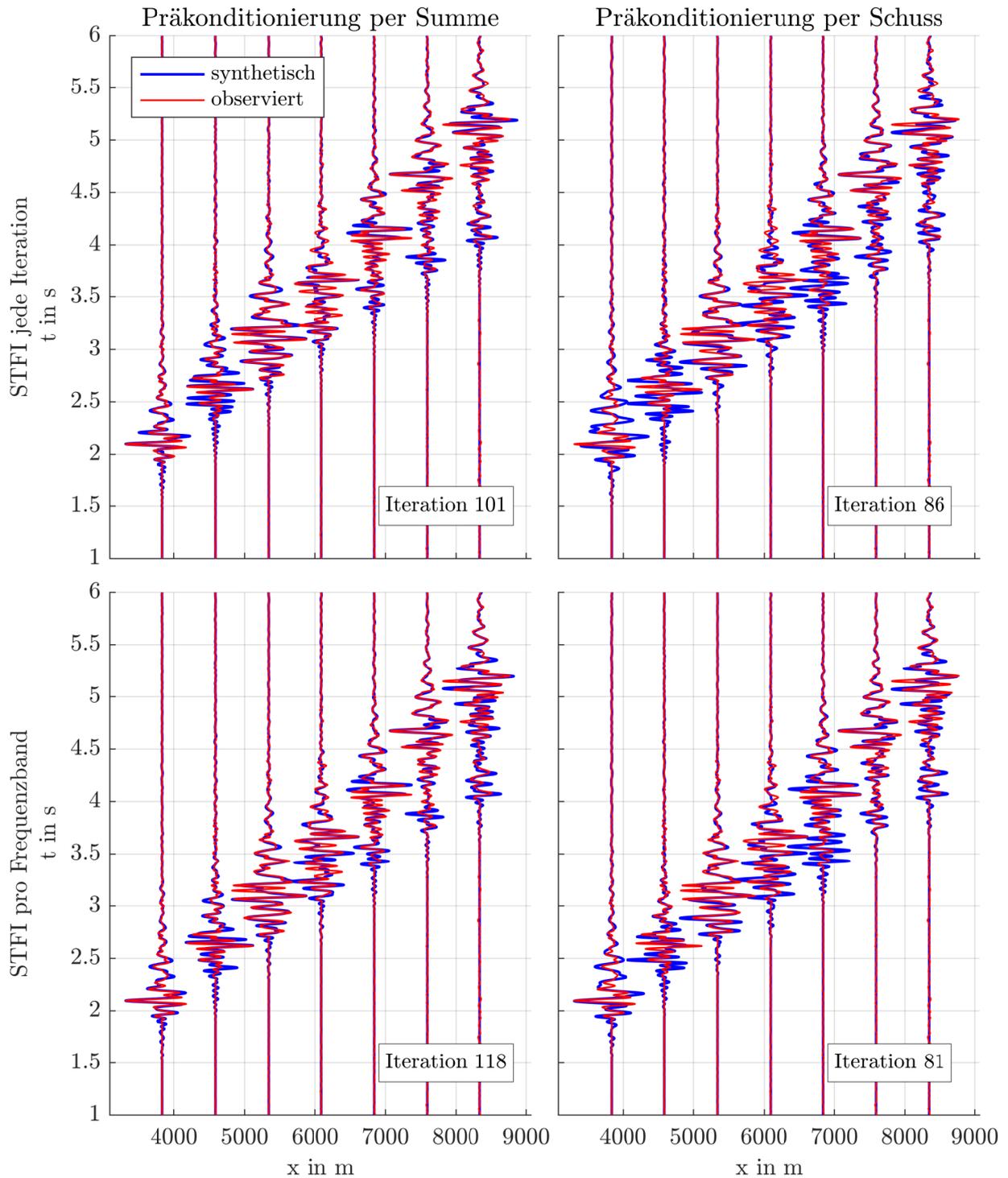


Abbildung 8.2.: Vergleich zwischen observierten und synthetischen Seismogrammen des 8. Schusses. Die Nummer der Iteration wurde so gewählt, dass jede Iteration am Abbruchkriterium des Frequenzbandes mit Grenzfrequenz von 16 Hz angelangt ist. Der zugehörige Misfit und die Frequenzbänder können in Abbildung 8.1 abgelesen werden.

8. Untersuchung von Quellsignalinversion und Prädiktionierung

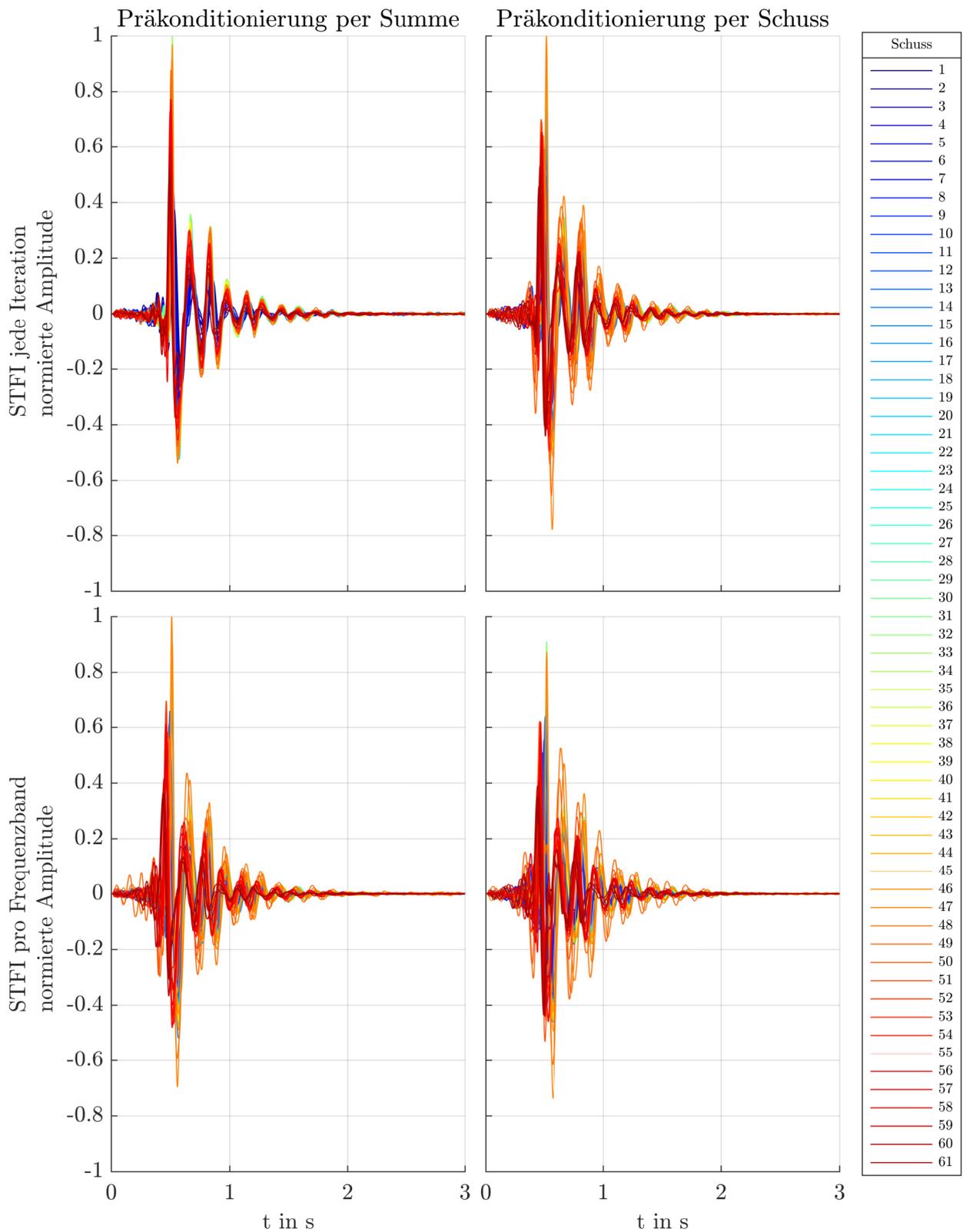
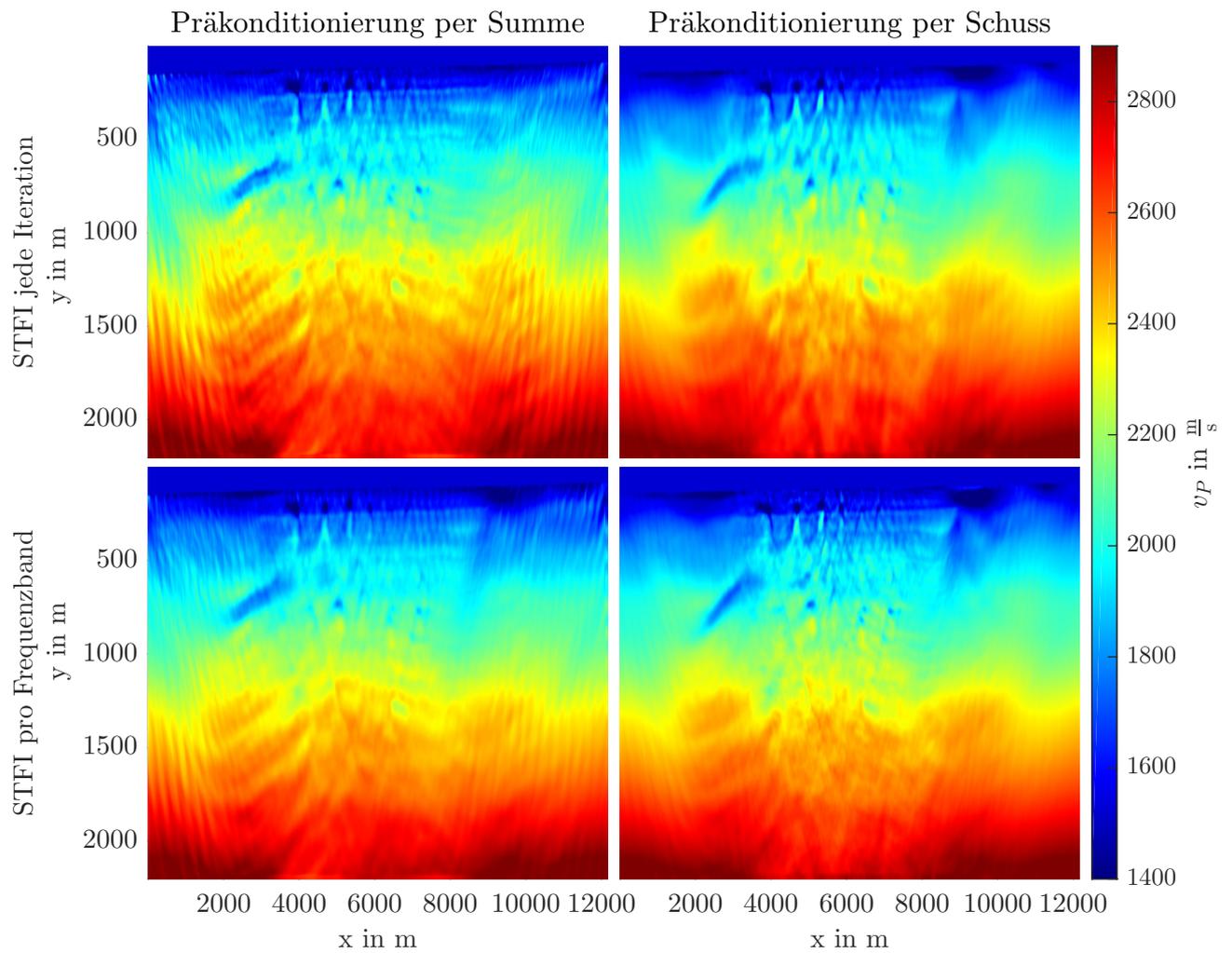


Abbildung 8.3.: Überlagerung der invertierten Quellsignale aller 61 Schüsse für jede der 4 Inversionen nach 160 Iterationen. Um ein Vergleich untereinander zu ermöglichen, wurde die Normierung aller 244 Signale mit einem einzigen Faktor vorgenommen.

Abbildung 8.4.: Invertierte v_P -Modelle nach 160 Iterationen.

9. Fazit

Mit dieser Arbeit konnte gezeigt werden, dass die Wellenforminversion in der Lage ist, geführte Wellen in marinen seismischen Daten zu verarbeiten. Meist werden diesen Wellen wenig Beachtung geschenkt, da sie keine Informationen von tieferen Schichten tragen. Allerdings verbessert ihre Verwendung die Rekonstruktion der oberflächennahen Bereiche, was folglich auch die Inversion der darunterliegenden Bereiche erleichtert. Besonders hervorzuheben ist die Anwendung einer minimalen Präprozessierung auf die Felddaten. Die notwendigen Manipulationen der 3D-2D Transformation und der Interpolation waren ausreichend. Zudem wurde auf eine Glättung der Gradienten verzichtet. Diese Reduzierung der Arbeitsschritte auf das Notwendige machen die Inversion leichter vergleichbar.

Da die observierten Daten aus einer realen Messung hervorgehen, ist das wahre v_p -Modell nicht bekannt. Für die Bewertung der Inversionsergebnisse war es somit wichtig alternative Kriterien zu finden. Die Evolution des Misfits, der Vergleich von Seismogrammspuren, Langsamkeits-Frequenz-Spektren, die invertierten Quellsignale und ein Vergleich mit einem Migrationsmodell ermöglichten eine Beurteilung der Inversion.

Im zweiten Teil der Arbeit wurde die Präkonditionierung des Gradienten mit der inversen Diagonalen der approximierten Hessematrix und die Wirkung der Quellsignalinversion untersucht. Hier konnten zwei Erkenntnisse gewonnen werden. Zum einen können durch die Präkonditionierung des Gradienten per Schuss, in dem hier gezeigten Beispiel, die Quellartefakte deutlich reduziert werden. Da dies ohne Erhöhung des Rechenaufwandes erreicht wird, sollte diese Variante auch in zukünftigen Wellenforminversionen getestet werden. Zum anderen führt die Quellsignalinversion in jeder Iteration nur zu einem leicht niedrigeren Misfit und dies auch nur bei der Kombination mit Präkonditionierung per Summe. Angesichts des hierfür notwendigen Mehraufwandes bezüglich der Rechenleistung, wird die Quellsignalinversion pro Frequenzband als ausreichend angesehen.

Literaturverzeichnis

- Aki, K. and Richards, P. (2002). *Quantitative Seismology*. University Science Books.
- Bohlen, T., De Nil, D., Köhn, D., and Jetschny, S. (2015). *Users Guide SOFI2D: Seismic modeling with finite differences*. Karlsruher Institut für Technology.
- Bradley, A. M. (2013). Pde-constrained optimization and the adjoint method.
- Bunks, C., Saleck, F. M., Zaleski, S., and Chavent, G. (1995). Multiscale seismic waveform inversion. *Geophysics*, 60(5):1457–1473.
- Drijkoningen, G. G. (2003). Seismic data acquisition ta 3600. Section Applied Geophysics Petrophysics, Delft University of Technology.
- Forbriger, T., Groos, L., and Schäfer, M. (2014). Line-source simulation for shallow-seismic data. part 1: Theoretical background. *Geophysical Journal International*, 198(3):1387–1404.
- Gardner, G., Gardner, L., and Gregory, A. (1974). Formation velocity and density - the diagnostic basics for stratigraphic traps. *Geophysics*, 39(6):770–780.
- Gerdtts, M. (2007). Skript, Optimierung. University of Birmingham, School of Mathematics.
- Gutenberg, B. (1914). Ueber erdbebenwellen. vii a. beobachtungen an registrierungen von fernbeben in göttingen und folgerung über die konstitution des erdkörpers (mit tafel). *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, 1914:125–176.
- Hadamard, J. (1902). Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton University Bulletin*, pages 49–52.

- Hestenes, M. R. and Stiefel, E. (1952). *Methods of conjugate gradients for solving linear systems*, volume 49. NBS.
- Kamath, N. and Tsvankin, I. (2016). Elastic full-waveform inversion for vti media: Methodology and sensitivity analysis. *Geophysics*, 81(2):C53–C68.
- Klein, G., Bohlen, T., Theilen, F., Kugler, S., and Forbriger, T. (2005). Acquisition and inversion of dispersive seismic waves in shallow marine environments. *Marine Geophysical Researches*, 26(2-4):287–315.
- Landau, L. and Lifschitz, E. (1997). *Lehrbuch der theoretischen Physik*. Harri Deutsch.
- Lehmann, I. (1936). P': Publications du bureau central séismologique international. *A14*, pages 87–115.
- Levander, A. (1988). Fourth-order finite-difference P-SV seismograms. *Geophysics*, 53(11):1425–1436.
- Liu, Q. and Tromp, J. (2006). Finite-frequency kernels based on adjoint methods. *Bulletin of the Seismological Society of America*, 96(6):2383–2397.
- McMechan, G. A. and Yedlin, M. J. (1981). Analysis of dispersive waves by wave field transformation. *Geophysics*, 46(6):869–874.
- Menke, W. (1989). *Geophysical Data Analysis: Discrete Inverse Theory*. International Geophysics. Elsevier Science.
- Neumann, M. (2013). Computational physics. Institut für Experimentalphysik der Universität Wien.
- Nocedal, J. and Wright, S. (2006). *Numerical Optimization*. Springer.
- Pekeris, C. L. (1948). Theory of propagation of explosive sound in shallow water. *Geological Society of America Memoirs*, 27:1–116.

- Pica, A., Diet, J., and Tarantola, A. (1990). Nonlinear inversion of seismic reflection data in a laterally invariant medium. *Geophysics*, 55(3):284–292.
- Plessix, R.-E. (2006). A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophysical Journal International*, 167(2):495.
- Pratt, R. G. (1999). Seismic waveform inversion in the frequency domain, part 1: Theory and verification in a physical scale model. *Geophysics*, 64(3):888–901.
- Press, F. and Ewing, M. (1950). Propagation of explosive sound in a liquid layer overlying a semi-infinite elastic solid. *Geophysics*, 15(3):426–446.
- Shearer, P. M. (2009). *Introduction to Seismology*. Cambridge University Press.
- Shewchuk, J. R. et al. (1994). An introduction to the conjugate gradient method without the agonizing pain.
- Shin, C., Yoon, K., Marfurt, K. J., Park, K., Yang, D., Lim, H. Y., Chung, S., and Shin, S. (2001). Efficient calculation of a partial-derivative wavefield using reciprocity for seismic imaging and inversion. *Geophysics*, 66(6):1856–1863.
- Tarantola, A. (1988). Theoretical background for the inversion of seismic waveforms including elasticity and attenuation. *Pure and Applied Geophysics*, 128(1):365–399.
- Virieux, J. (1986). P-SV wave propagation in heterogeneous media: velocity-stress finite-difference method. *Geophysics*, 51(4):889–901.

A. Videos

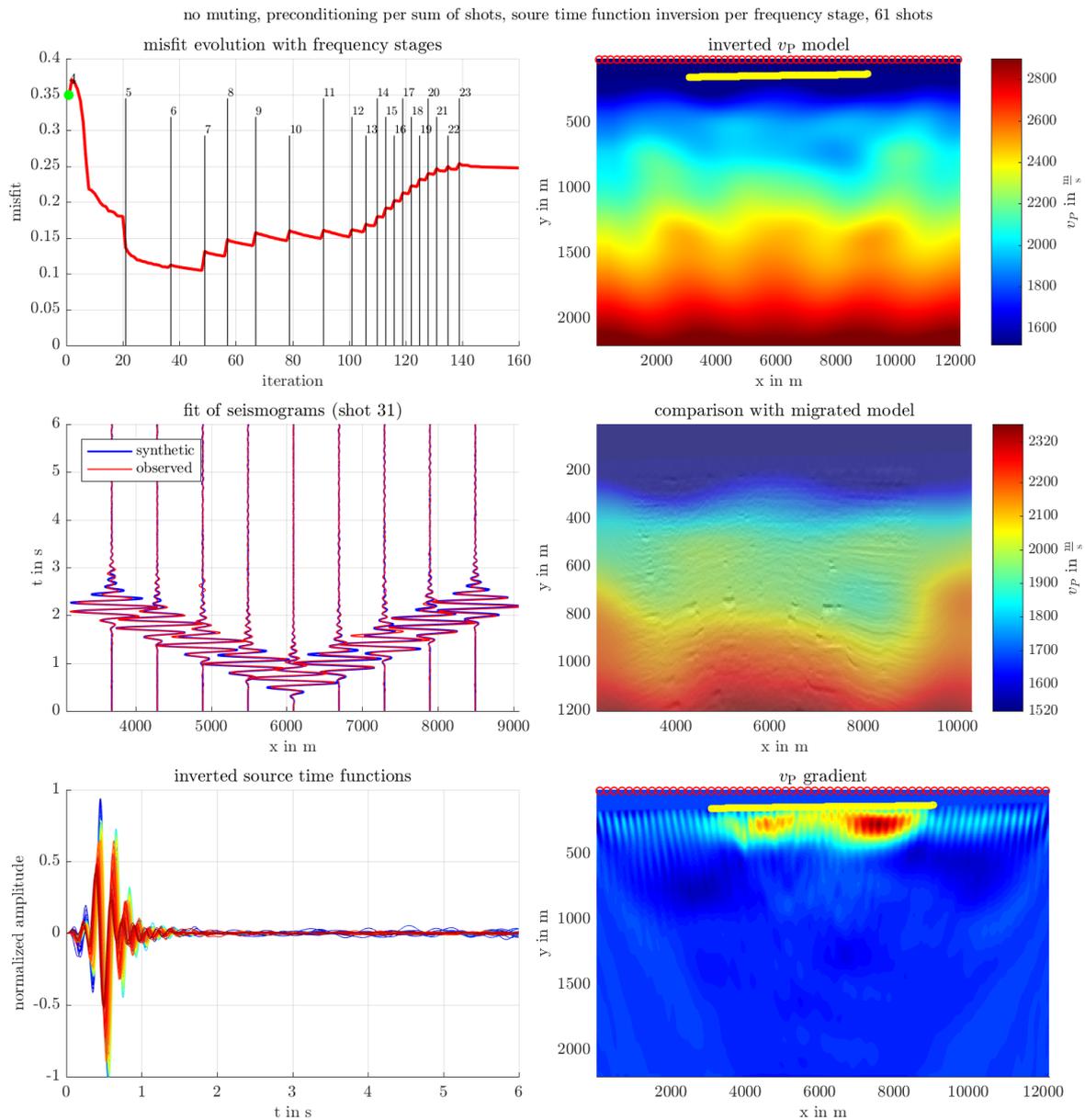


Abbildung A.1.: Evolution verschiedener Inversionsergebnisse der Referenzlösung (vgl. Kapitel 7): Misfit, v_P -Modell, Vergleich zwischen synthetischen und beobachteten Seismogrammen, Überlagerung von Migrationsergebnis und v_P -Modell, invertierte Quellsignale aller Schüsse und v_P -Gradient.

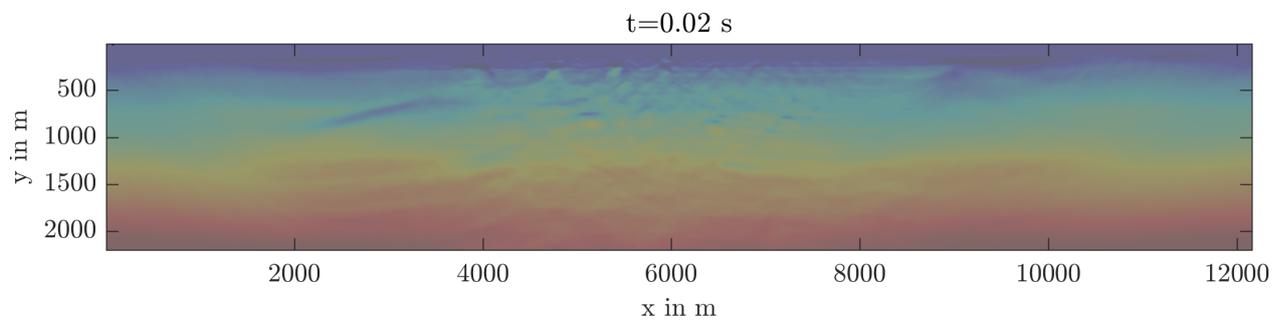


Abbildung A.2.: Simulation der akustischen Wellenausbreitung im finalen Modell (160. Iteration, Einstellung: STFI pro Frequenzband, Prakonditionierung per Schuss) von Schuss 1. Zur besseren Sichtbarkeit der Ausbreitungswege sind die Amplituden normiert. Gut zu Erkennen sind Tauchwellen, welche Informationen von tieferen Schichten sammeln und die multiplen Reflexionen zwischen Oberflache und Meeresboden, die zur Ausbildung der gefuhrten dispersiven Wellen fuhren.