

Velocity model building using analytic and model-based diffraction traveltime functions

Geschwindigkeitsmodellbildung unter Benutzung analytischer und modellbasierter Diffraktionslaufzeitfunktionen

Zur Erlangung des akademischen Grades eines
DOKTORS DER NATURWISSENSCHAFTEN

von der Fakultät für Physik der
Universität Karlsruhe (TH)
genehmigte

DISSERTATION

von

Dipl.-Geophys. Tilman Klüver

aus

Lüneburg

Tag der mündlichen Prüfung:

01. Juni 2007

Referent:

Prof. Dr. Peter Hubral

Korreferent:

Prof. Dr. Dirk Gajewski

Abstract

The determination of a structural image of the subsurface from seismic data by depth migration requires a seismic velocity model. Commonly used methods for the determination of such velocity models in laterally inhomogeneous media are, among many others, reflection traveltome tomography and migration-based velocity analysis. These methods have several drawbacks like the tremendous amount of picking in the seismic prestack data required for the former method and the repeated application of prestack depth migration necessary in the latter class of methods. In this thesis, two tomographic inversion methods for the determination of smooth isotropic inhomogeneous velocity models are discussed which partially overcome these drawbacks. Both methods make use of information related to diffraction traveltome operators.

The first method uses attributes parameterising second-order approximations of true diffraction traveltome functions. These attributes are coefficients in second-order traveltome approximations in the midpoint and offset dimension of the reflection events in the seismic data. They may be automatically extracted from the seismic prestack data by means of coherence analyses. In this thesis, the common-reflection-surface (CRS) stack method is used for this purpose. The identification of reliable pick locations in the zero-offset (ZO) section/volume simulated by the CRS stack method is much easier than picking in the prestack data due to the increased signal-to-noise ratio. As each pick is associated with an isolated reflection point, the pick locations do not need to follow continuous, interpreted reflection events in the ZO stack section/volume. There is even no need for the existence of such continuous horizons, the identification of locally coherent events is sufficient. The inversion process aims at minimising the misfit between the extracted attributes and their forward-modelled counterparts. As the attributes only approximate the true diffraction traveltome functions up to second order, the resulting velocity models are considered only to be consistent with the seismic prestack data up to second order as well.

A new tomographic inversion scheme is introduced which is suitable to update these models and overcomes the limitation to second order traveltome approximations. It aims at ensuring the kinematic equivalence between common-reflection-point (CRP) gathers and common-image-gathers (CIG) for all reflection points used in the inversion process. For this purpose, model-based diffraction traveltome functions are used for a moveout correction of the prestack data. This moveout correction results in a small prestack data volume for each pick in which the traces are sorted according to the illumination and scattering angle of the seismic energy at the reflection point. In these moveout-corrected prestack data volumes time residuals and time dip residuals with respect to the illumination angle are extracted along the model-based CRP trajectory. The CRP trajectory is that trajectory in the seismic prestack data along which all signals stemming from the same reflection point are found. The minimisation of both types of residuals ensures that the CRP trajectory is actually a trajectory of stationary points, i. e., a trajectory along which the diffraction traveltome operator is tangent to the reflection event in

the seismic data. To ensure that the time and time dip residuals are actually minimised along the CRP trajectories related to the input pick locations, the inversion scheme additionally tries to minimise the misfit between the extracted and forward-modelled attributes describing the ZO stationary point. In contrast to Kirchhoff prestack depth migration the diffraction traveltime operator only needs to be calculated for a relatively small number of selected reflection points. This makes the used move-out correction much faster than the computation of CIGs in conventional migration-based velocity analysis methods.

In this thesis, the theory for both inversion schemes as well as several applications to synthetic and real seismic data are presented. The inversion scheme based on second-order approximations of diffraction traveltime functions is discussed with special emphasis placed on the pre-conditioning and automatic extraction of the input data. Several practical and implementation-related aspects are described and demonstrated using 2D and 3D synthetic and real data examples. The inversion scheme using model-based diffraction traveltime operators is introduced and theoretical as well as practical aspects are discussed for the 2D case. The method is demonstrated using a 2D synthetic data example. As both inversion schemes are based on the same model description and partially use identical input data quantities they can be successively applied in a velocity model building work flow: it starts with the highly automated inversion scheme based on approximate diffraction traveltime functions followed by an update using the presented new inversion scheme which overcomes the limitation to second order. This consistent velocity model building work flow is demonstrated using a complex 2D synthetic data example.

Zusammenfassung

Vorbemerkung: Die vorliegende Dissertation ist bis auf diese Zusammenfassung in englischer Sprache geschrieben. Da auch in der deutschen Sprache einige englische Fachausdrücke gebräuchlich sind, wurde bei diesen Ausdrücken auf eine Übersetzung verzichtet. Sie werden, mit Ausnahme ihrer groß geschriebenen Abkürzungen, *kursiv* dargestellt.

Einführung

Die Reflexionsseismik spielt eine zentrale Rolle in der Exploration nach Kohlenwasserstoffreserven. Reflexionsseismische Datenakquisition wird sowohl auf See wie auch auf der Landoberfläche durchgeführt. In der reflexionsseismischen Methode wird mit Hilfe kontrollierter Quellen seismische Energie in den Untergrund abgestrahlt. Diese Energie propagiert durch den Untergrund und erreicht teilweise wieder die Messoberfläche, nachdem sie an Diskontinuitäten der elastischen Eigenschaften des Untergrundes reflektiert wurde. Die resultierenden Partikelverschiebungen bzw. Druckveränderungen werden von den Empfängern als diskrete Zeitreihen aufgezeichnet. Sowohl die Laufzeiten als auch die Amplituden der gemessenen Signale enthalten vielfältige Informationen über den Untergrund, die es aus den Messdaten zu extrahieren gilt.

Ein reflexionsseismisches Einzelexperiment besteht in der Regel aus einer Quelle, die die Energie abstrahlt, und einer Vielzahl von Empfängern mit verschiedenem Abstand (*offset*) zur Quelle, die die Signale registrieren. Jede dieser Zeitreihen stellt eine sogenannte seismische Spur dar. Durch die Wiederholung eines solchen Einzelexperiments mit jeweils anderer Konfiguration von Quelle und Empfängern entsteht ein mehrfach überdeckter seismischer Datensatz. Dieser enthält redundante Informationen über die zu untersuchende Struktur im Untergrund. Bei einer klassischen zweidimensionalen Datenakquisition sind Quellen und Empfänger längs einer Linie angeordnet, während bei einer dreidimensionalen Datenakquisition Quellen und Empfänger beliebig auf der Messoberfläche verteilt sein können.

Die Redundanz in den mehrfach überdeckten, seismischen Daten spielt eine wichtige Rolle für deren weitere Bearbeitung und Auswertung. Erste Bearbeitungsschritte zielen darauf ab, die in der Reflexionsseismik relevanten Signale (Primärreflexionen) hervorzuheben und nicht nutzbare Signale, etwa multiple Reflexionen und Rauschen, zu unterdrücken. Zu diesem Zweck wird häufig eine Summation (Stapelung, *stack*) von Signalen mit verschiedenem Abstand (*offset*) aber gleichem Mittelpunkt (*midpoint*) zwischen Quelle und Empfänger durchgeführt. Dieser Schritt reduziert das Datenvolumen und erhöht das Signal/Rauschen (S/R) Verhältnis. Der resultierende Datensatz simuliert das Ergebnis einer seismischen Messung mit koinzidenten Quellen und Empfängern (*zero-offset*).

Ein Hauptziel der Reflexionsseismik ist es, aus den gemessenen Daten ein detailliertes strukturelles Abbild des Untergrunds zu erstellen. Dafür wird eine sogenannte Tiefenmigration durchgeführt, deren Ziel es ist, die Wellenausbreitungseffekte im Untergrund rückgängig zu machen. Dadurch werden die Reflexionsereignisse in den Messdaten in Reflektorabbilder überführt. Zur Durchführung der Tiefenmigration wird ein Modell der Ausbreitungsgeschwindigkeit seismischer Wellen im Untergrund benötigt. Dieses Modell ist im allgemeinen unbekannt und muß unter Zuhilfenahme weiterer Informationen aus den seismischen Messdaten gewonnen werden. Die Daten selbst enthalten generell nicht genügend Informationen um die tatsächliche Verteilung der Ausbreitungsgeschwindigkeiten seismischer Wellen zu bestimmen. Es lässt sich normalerweise aber ein Modell bestimmen, das konsistent mit den gemessenen Daten und eventuell vorhandener Zusatzinformation ist.

Die Erstellung eines solchen Modells ist mit einem nichtlinearen Inversionsproblem verknüpft, das im allgemeinen iterativ gelöst wird. Dafür sind verschiedene Verfahren in Gebrauch. Die meisten basieren auf Kriterien, mit denen die Konsistenz des Modells mit den gemessenen Daten überprüft werden kann. Sie unterscheiden sich darin, wie diese Kriterien definiert sind, wie Inkonsistenzen in Veränderungen des Modells übersetzt werden und wie das Modell parametrisiert ist. Ein übliches Verfahren ist die sogenannte Laufzeittomographie. In diesem Verfahren wird ein Modell, bestehend aus Reflektoren und Ausbreitungsgeschwindigkeiten in den Schichten dazwischen, so lange gezielt verändert, bis die Abweichung von in diesem Modell modellierten Laufzeiten zu Laufzeiten in den Messdaten minimiert ist. Üblicherweise werden die zu modellierenden Laufzeiten entlang kohärenter Reflexionsereignisse in den seismischen Daten ausgewählt. Dies ist aufgrund der großen Datenmenge sehr schwierig und zeitaufwendig. In seismischen Daten mit schlechtem S/R Verhältnis lassen sich kohärente Reflexionsereignisse mitunter kaum erkennen. Andere Verfahren zur Bestimmung eines Geschwindigkeitsmodells beruhen auf der wiederholten Anwendung von Tiefenmigration. In einem konsistenten Modell müssen alle migrierten Sektionen, die jeweils Daten mit gleichem Abstand zwischen Quelle und Empfänger enthalten, kinematisch identisch sein. Abweichungen von diesem Kriterium werden gezielt in Veränderungen des Modells übersetzt. Die wiederholte Anwendung der Tiefenmigration ist allerdings extrem zeitaufwendig und daher ein Nachteil dieser Vorgehensweise.

In dieser Arbeit werden zwei Verfahren diskutiert, die diese Nachteile teilweise umgehen. Beide Verfahren nutzen Eigenschaften von Diffraktionslaufzeitfunktionen, d. h. Eigenschaften der kinematischen Antwort einer Punktquelle im Untergrund. Während das erste Verfahren Parameter benutzt, welche die wahren Diffraktionslaufzeitfunktionen bis zur zweiten Ordnung approximieren, verwendet das zweite Verfahren modellierte Diffraktionslaufzeitfunktionen. Es macht daher keine Einschränkungen an die Komplexität des Geschwindigkeitsmodells. Da beide Verfahren mit kinematischen Informationen arbeiten, die von isolierten Tiefenpunkten stammen, müssen keine Reflektoren in die Modellparametrisierung aufgenommen werden. Jeder Datenpunkt kann als unabhängig von allen anderen betrachtet werden. Daher müssen keine zusammenhängenden Reflexionsereignisse in den Daten verfolgt werden. Die Datenpunkte werden in gestapelten, simulierten *zero-offset* Sektionen auf lokal kohärenten Ereignissen ausgewählt. Die Stapelsektionen werden mit Hilfe der *common-reflection-surface* (CRS) *stack* Methode erstellt und weisen, verglichen mit den Messdaten, ein stark verbessertes S/R Verhältnis bei gleichzeitig erheblich reduzierter Datenmenge auf. Dadurch wird die Auswahl der für die tomographische Inversion geeigneten Datenpunkte erheblich vereinfacht und kann teilweise automatisiert werden.

Seismische Strahlentheorie

Beide vorgestellte Verfahren zur Bestimmung von seismischen Geschwindigkeitsmodellen machen Gebrauch von verschiedenen Elementen der seismischen Strahlentheorie. Da im Rahmen dieser Arbeit nur kinematische Aspekte der seismischen Wellen eine Rolle spielen, wird die Beschreibung der Strahlenmethode hierauf beschränkt.

Die Strahlentheorie entsteht aus einer approximativen Lösung der elastodynamischen Wellengleichung, die nur für hochfrequente Wellen gültig ist. Hochfrequent bedeutet, dass die Eigenschaften des von der Welle durchlaufenen Mediums sich nur auf Längenskalen ändern, die deutlich größer sind als die dominierenden Wellenlängen der zu beschreibenden Welle. Im Rahmen der Hochfrequenzapproximation werden die kinematischen Aspekte von der Eikonalgleichung beschrieben. Dies ist eine nichtlineare, partielle Differentialgleichung erster Ordnung, die die Ausbreitung von Wellenfronten beschreibt. Die Eikonalgleichung lässt sich mit der Methode der Charakteristiken lösen, die zu einem System von gekoppelten, linearen Differentialgleichungen erster Ordnung für die Ortskoordinaten und des zugehörigen *slowness*-Vektors führt. Die Charakteristiken sind Lösungen dieses Systems. Die Projektion einer Charakteristik in den dreidimensionalen Ortsraum wird als Strahl bezeichnet, das Gleichungssystem als *ray tracing* System. Im in dieser Arbeit ausschließlich betrachteten Fall von Wellenausbreitung in isotropen Medien stehen die Strahlen senkrecht auf den Wellenfronten.

Besondere Bedeutung kommt in dieser Arbeit dem paraxialen *ray tracing* System zu. Es beschreibt Strahlen in der unmittelbaren Umgebung eines Referenzstrahls mit Hilfe von Größen, die auf dem Referenzstrahl selbst berechnet werden. Das paraxiale *ray tracing* System ist ein System gewöhnlicher Differentialgleichungen erster Ordnung, dessen Lösungen sich mit Hilfe einer sogenannten Propagatormatrix darstellen lassen. Formal identisch ist das sogenannte dynamische *ray tracing* System. Es erlaubt u. a. die Approximation von Laufzeiten vorgegebener Wellen in der Nähe des Referenzstrahls bis zur zweiten Ordnung. Eine solche Laufzeitapproximation einer Welle, die von einer Punktquelle ausgeht, wird in dieser Arbeit genauer betrachtet.

Um Abweichungen der modellierten Daten von den gemessenen innerhalb der tomographischen Inversionsverfahren mit Änderungen des Modells zu verknüpfen, werden des Weiteren einige Ergebnisse der Strahlen-Perturbationstheorie benutzt. Sie ermöglichen die Auswirkungen von Störungen der Anfangsbedingungen als auch von Störungen des Geschwindigkeitsmodells auf den Strahlverlauf bis zur ersten Ordnung zu quantifizieren. Auch die Auswirkungen dieser Störungen auf die Propagatormatrix lassen sich mit Hilfe der Strahlen-Perturbationstheorie untersuchen.

Modellbildung mit analytischen Diffraktionslaufzeitfunktionen

Das erste, hier diskutierte Verfahren zur Bestimmung von Geschwindigkeitsmodellen wurde von [Duvencek \(2004a\)](#) vorgestellt. In dieser Arbeit liegt der Schwerpunkt auf einigen Erweiterungen sowie auf der Automatisierung dieses Verfahrens.

Im tomographischen Inversionsverfahren werden Parameter benutzt, die Koeffizienten in einer Approximation zweiter Ordnung bzgl. der *midpoint*- und *offset*-Koordinaten von Diffraktionslaufzeitfunktionen sind. Diese Parameter lassen sich mit Hilfe der CRS Stapelmethode aus den seismischen Daten extrahieren. Die Parameter, die den CRS Stapeloperator beschreiben, lassen sich physikalisch mit zwei hypothetischen Wellenfronten erklären, die durch ein explodierendes Reflektorsegment bzw.

durch eine Punktquelle auf einem Reflektor entstehen. Letztere Wellenfront wird als NIP-Wellenfront bezeichnet. Der Reflexionspunkt des zugehörigen *zero-offset* Strahls, der im hier betrachteten Fall mit dem Normalstrahl übereinstimmt, wird als *normal-incidence-point* (NIP) bezeichnet. Die Größen, welche die kinematischen Eigenschaften der NIP-Welle an der Messoberfläche erklären, sind die Eingangsdaten für das tomographische Inversionsverfahren. Sie werden an ausgewählten Lokationen aus den CRS Stapelergebnissen extrahiert.

Das Inversionsverfahren beruht auf folgender Abbildungsbedingung: Werden die auftauchenden NIP-Wellen entlang der Normalstrahlen zurück in den Untergrund propagiert, so müssen sie bei Laufzeit Null am NIP fokussieren. Diese Bedingung ist äquivalent zur Fokussierung aller Energie, die entlang eines bis zur zweiten Ordnung approximierten Diffraktionslaufzeitoperators in den seismischen Daten gefunden wird, an einem gemeinsamen Abbildungspunkt mit Hilfe einer Tiefenmigration.

Jeder eine NIP-Welle repräsentierende Datenpunkt besteht aus der Einweglaufzeit der Welle entlang des Normalstrahls, dessen Auftauchpunkt, sowie den ersten und zweiten horizontalen Laufzeitableitungen der NIP-Welle an dieser Stelle. Der Auftauchpunkt ist identisch mit der Lokation in der simulierten CRS Stapelsektion, die Einweglaufzeit der NIP-Welle ist durch die halbe Zeit in der CRS Stapelsektion gegeben. Das Modell besteht aus der Geschwindigkeitsverteilung im Untergrund sowie aus den NIP-Lokationen und der Orientierung der Normalstrahlen am zugehörigen NIP. Das Geschwindigkeitsmodell wird mit B-Splines dargestellt. Dies sind glatte, lokalisierte Spline-Basisfunktionen. Die Koeffizienten dieser Spline-Darstellung sind die eigentlichen Komponenten des Geschwindigkeitsmodells.

In der Implementierung werden die NIP-Wellen allerdings nicht zurück in den Untergrund propagiert. Stattdessen werden zu jedem Datenpunkt die entsprechenden Größen modelliert und mit den Eingangsdaten verglichen. Die Modellierung wird mit Hilfe des dynamischen *ray tracing* durchgeführt. Durch tomographische Inversion wird ein Modell bestimmt, in dem der Unterschied zwischen den Eingangsdaten und den entsprechenden modellierten Größen minimiert wird. Dieses nichtlineare Optimierungsproblem wird iterativ im Sinne der kleinsten Fehlerquadrate gelöst. Die dafür benötigten Fréchet-Ableitungen werden mit Hilfe der Strahlen-Perturbationstheorie berechnet. Da die Eingangsdaten in der Regel keine ausreichenden Informationen enthalten, um ein eindeutiges Inversionsergebnis zu erhalten, wird das Inversionsproblem durch die Minimierung der zweiten Ableitungen des Geschwindigkeitsmodells regularisiert. Außerdem können noch weitere Bedingungen an das Modell gestellt werden, um zum Beispiel vorhandene geologische Kenntnisse zu berücksichtigen.

Ein Schwerpunkt dieser Arbeit ist die Entwicklung von Verfahren, die die automatische Extraktion der Eingangsdaten für die Tomographie aus den CRS Stapelergebnissen ermöglichen. Besonderer Wert wird dabei auf das physikalisch gerechtfertigte Glätten der Parameter gelegt, die den CRS Stapeloperator beschreiben. Dies geschieht durch Filterungsprozesse, die lokal in einem kleinen Fenster angewendet werden, das entlang des betrachteten Reflexionsereignisses ausgerichtet ist. Mit Hilfe einer robusten, linearen Regressionstechnik lassen sich des Weiteren Fehlergrenzen für die Eingangsdaten direkt aus den Streuungen der CRS Stapelparameter bestimmen. Diese Fehlerabschätzungen werden in der Tomographie als Kovarianzoperatoren benutzt und ermöglichen somit eine Gewichtung der einzelnen Datenkomponenten, die die Ungenauigkeit in der Bestimmung der CRS Stapelparameter berücksichtigt. Dieser Aspekt der Arbeit wird anhand eines synthetischen 2D Datenbeispiels demonstriert.

Die ursprüngliche Implementierung für den 3D Fall wird dahingehend erweitert, die Azimutabhängigkeit der zweiten Laufzeitableitungen zu berücksichtigen. Vorausgesetzt, die Akquisitionsgeometrie lässt die Bestimmung dieser Azimutabhängigkeit zu, erlaubt diese Erweiterung die größtmögliche

Komplexität (innerhalb der zweiten Ordnung) des CRS Stapeloperators in der Tomographie zu berücksichtigen.

Da das hier beschriebene tomographische Inversionsverfahren auf einer Approximation zweiter Ordnung der wahren Diffraktionslaufzeitfunktionen basiert, können in gewissem Sinne die erhaltenen Geschwindigkeitsmodelle auch nur bis zur zweiten Ordnung konsistent mit den seismischen Daten sein. Sinnvolle Ergebnisse können nur erhalten werden, wenn die Eingangsdaten die kinematischen Eigenschaften von NIP-Wellen hinlänglich genau beschreiben.

Modellbildung mit modellbasierten Diffraktionslaufzeitfunktionen

Um die Beschränkung der oben beschriebenen Methode auf Approximationen zweiter Ordnung aufzuheben, stelle ich in dieser Arbeit ein neues tomographisches Inversionsverfahren vor. Es nutzt modellbasierte Diffraktionslaufzeitfunktionen und macht keine Einschränkungen an die Komplexität des Modells. Die Diffraktionslaufzeitfunktionen werden direkt für jeden Eingangsdatenpunkt berechnet. Dies ist wesentlich effizienter als die Durchführung einer kompletten Tiefenmigration, wie sie in migrationsbasierten Verfahren zur Geschwindigkeitsmodellbildung notwendig ist. Die theoretischen und praktischen Aspekte des Verfahrens werden für den 2D Fall diskutiert.

Das Verfahren nutzt die gleiche Modellbeschreibung, die auch für die oben vorgestellte Inversionsmethode benutzt wird. Jeder Datenpunkt besteht aus der Einweglaufzeit entlang eines Normalstrahls, dessen Auftauchlokation, sowie der horizontalen *slowness* am Auftauchpunkt. Diese Datenpunkte werden mit den gleichen Techniken in den CRS Stapelergebnissen ausgewählt wie für das andere, bereits vorgestellte Inversionsverfahren. Jedem dieser Datenpunkte ist in der Tiefe wieder ein Reflexionspunkt zugeordnet. Die zweiten Laufzeitableitungen der NIP-Wellenfront werden durch Zeit- und Neigungsabweichungen des modellierten Diffraktionslaufzeitoperators vom Reflexionsereignis in den seismischen Daten entlang der sogenannten *common-reflection-point* (CRP) Trajektorie ersetzt. Entlang der CRP Trajektorie befinden sich in den seismischen Daten alle Signale von Wellen, die am gleichen Tiefenpunkt reflektiert wurden.

Um die Zeit- und Neigungsabweichungen für jeden Datenpunkt zu bestimmen, wird ein Diffraktionslaufzeitoperator für jeden zugehörigen Tiefenpunkt durch *ray tracing* modelliert. Alle Daten innerhalb eines bestimmten räumlichen Bereichs um die Auftauchlokation des Normalstrahls werden dann zeitlich relativ zu den modellierten Diffraktionslaufzeiten betrachtet. Der Diffraktionslaufzeitoperator bestimmt also die Fläche mit Zeit Null. Gleichzeitig werden die Daten nach dem Beleuchtungs- und Streuwinkel am Tiefenpunkt sortiert. Auf diese Weise entsteht für jeden Datenpunkt ein kleines, zeitlich korrigiertes, seismisches Datenvolumen. Der Zusammenhang zwischen Quell- und Empfängerlokationen an der Oberfläche und den Winkeln am Reflexionspunkt ist durch die Modellierung des Diffraktionslaufzeitoperators bekannt.

Wurde für die Modellierung der Diffraktionslaufzeitfunktionen ein Geschwindigkeitsmodell benutzt, das konsistent mit den seismischen Daten ist, befinden sich alle Signale entlang der CRP Trajektorie in diesen Datenvolumen bei Zeit Null und dem Beleuchtungswinkel, der mit dem Reflektorneigungswinkel übereinstimmt. Zeitliche Abweichungen lassen sich durch eine Kreuzkorrelation mit einem Teil der zu diesem Datenpunkt gehörigen *zero-offset* Spur bestimmen. Für ein konsistentes Modell müssen sich die Signale allerdings nicht nur bei Zeit Null befinden, sondern auch die Neigung des Reflexionsereignisses bzgl. des Beleuchtungswinkels in den zeitlich korrigierten Datenvolumen muß

entlang der CRP Trajektorie verschwinden. Dies garantiert die Tangentialität des Diffraktionslaufzeitoperators und des Reflexionsereignisses in den seismischen Daten. Abweichungen dieser Neigung lassen sich entlang der CRP Trajektorie für jeden Streuwinkel mit Hilfe der schon bekannten Zeitfehler durch Kohärenzanalysen bestimmen. Die Auswertung des Neigungskriteriums wird dadurch notwendig, dass isolierte Tiefenpunkte betrachtet werden, die mitunter große laterale Abstände haben. In einer lateral kontinuierlich durchgeführten, migrationsbasierten Geschwindigkeitsanalyse ist die Auswertung dieses Kriteriums implizit enthalten.

Das hier vorgestellte Inversionsverfahren hat zum Ziel, ein Geschwindigkeitsmodell zu bestimmen, in dem sowohl die Abweichung der *zero-offset* Eingangsdaten von den entsprechenden modellierten Größen als auch die extrahierten Zeit- und Zeitneigungsfehler entlang der CRP-Trajektorien minimiert werden. Dieses Inversionsproblem ist nichtlinear und wird, wie im oben beschriebenen, anderen Verfahren, iterativ gelöst. In jeder Iteration müssen dabei die Zeit- und Neigungsfehler bestimmt werden. Dieses Inversionsverfahren ist daher wesentlich aufwendiger als das oben beschriebene. Die Art der Regularisierung und Einbindung von zusätzlichen Bedingungen an das Modell wurden vom oben beschriebenen Verfahren übernommen. Dies, zusammen mit derselben Modellbeschreibung und teilweise identischen Eingangsdaten, macht dieses neue Inversionsverfahren zu einer idealen Methode, um die auf einer Approximationen der wahren Diffraktionslaufzeitfunktionen basierenden Modelle zu verbessern. Das neue tomographische Inversionsverfahren wird an einem synthetischen Datenbeispiel demonstriert.

Anwendungen

Als Abschluß der Arbeit stelle ich zwei Datenbeispiele vor, die diverse Aspekte der vorgestellten Inversionsverfahren noch einmal beleuchten. Im ersten Beispiel demonstriere ich den kompletten Prozess der Geschwindigkeitsmodellbildung an einem komplexen 2D synthetischen Datenbeispiel. Nach der Anwendung der CRS Stapelungsmethode werden automatisch geeignete Lokationen für die Extraktion der Eingangsdaten bestimmt. Dann wird ein Modell mit dem ersten der hier vorgestellten Verfahren bestimmt. Diese Modell weist nur ein geringes Maß an Details auf und führt nicht zu einem befriedigenden, strukturellen Abbild durch Tiefenmigration. Anschließend werden die Verbesserungen dieses Modell durch die Anwendung des zweiten in dieser Arbeit diskutierten Inversionsverfahrens verdeutlicht.

Im zweiten Beispiel wird ein realer seismischer 3D Datensatz betrachtet. Die Akquisitionsgeometrie erlaubt die Bestimmung von azimutabhängigen zweiten Ableitungen der Laufzeit. Nach der Anwendung der CRS Stapelungsmethode wird das erste der in dieser Arbeit diskutierten Inversionsverfahren mit zwei Eingangsdatensätzen angewendet. Im ersten Fall wird nur eine Komponente der Matrix der zweiten Laufzeitableitungen benutzt, während im zweiten Fall alle Matrixkomponenten eingehen. Die Unterschiede in den erhaltenen Modellen werden diskutiert und strukturelle Abbilder durch 3D Tiefenmigration erstellt.

Schlussfolgerungen und Ausblick

In dieser Arbeit wurden zwei tomographische Inversionsverfahren zur Bestimmung von Geschwindigkeitsmodellen für die Tiefenmigration diskutiert. Beide Verfahren benutzen Diffraktionslaufzeitfunktionen. Im ersten Verfahren werden die Eingangsdaten durch Größen gebildet, die die wahren

Diffraktionslaufzeitoperatoren bis zur zweiten Ordnung approximieren. Im zweiten Verfahren werden modellierte Diffraktionslaufzeitfunktionen benutzt, um die seismischen Daten zeitlich zu verschieben und Kriterien auszuwerten, die die Tangentialität der Diffraktionslaufzeitoperatoren und der Reflexionsereignisse messen. Die Lokationen der Eingangsdaten werden mit automatischen Selektierungsverfahren in den Ergebnissen der CRS Stapelungsmethode ausgewählt. In dieser Arbeit wurden dafür geeignete Strategien vorgestellt. Ebenfalls präsentiert wurden Verfahren für die Bestimmung von Fehlergrenzen der CRS Parameter. Die Anwendung dieser Fehlerabschätzungen als Kovarianzoperatoren in der Tomographie wurde diskutiert. Nacheinander angewendet ergeben beide Verfahren eine in sich konsistente Methode der Geschwindigkeitsmodellbestimmung. Durch das Betrachten isolierter Tiefenpunkte wurde die Bestimmung der Eingangsdaten gegenüber klassischen Verfahren wie der Reflexionslaufzeit-tomographie stark vereinfacht. Für beide Verfahren wurden theoretische wie auch praktische Aspekte beschrieben und anhand geeigneter Datenbeispiele verdeutlicht.

In zukünftigen Arbeiten, die sich mit den hier vorgestellten Methoden beschäftigen, sollte das Hauptaugenmerk auf einer weiteren Stabilisierung der Verfahren liegen. Dies betrifft insbesondere die Erkennung von Ausreißern in den automatisch extrahierten Eingangsdaten und die Regularisierung des Inversionsproblems. Letzteres könnte zum Beispiel durch einen Wechsel des Bereichs erreicht werden, in dem die Lokationen der Eingangsdaten gewählt werden. Möglichkeiten dafür werden durch die Verwendung von zeit- bzw. tiefenmigrierten CRS Stapelsektionen eröffnet.

Contents

Abstract	i
Zusammenfassung	iii
1 Introduction	1
1.1 Seismic data acquisition	1
1.2 Seismic data processing	2
1.3 Seismic imaging and velocities	3
1.3.1 Time migration	4
1.3.2 Depth imaging	4
1.4 Methods for velocity model estimation	4
1.4.1 Migration-based methods	5
1.4.2 Traveltime inversion	6
1.5 Migration velocity model building from diffraction traveltime operators	7
1.6 Structure of the thesis	8
1.7 Notation	9
2 Ray and ray perturbation theory	11
2.1 The elastodynamic wave equation	11
2.2 The solution of the eikonal equation	13
2.3 The reduced ray tracing system	14
2.4 Paraxial ray tracing	14
2.4.1 The ray propagator matrix	15
2.5 Dynamic ray tracing	17
2.6 Second-order traveltime approximation	17

2.7	Ray perturbation theory	18
2.8	Ray-centred coordinates	20
2.8.1	Transformation to Cartesian coordinates	22
2.8.2	Paraxial ray tracing in ray-centred coordinates	22
2.8.3	Dynamic ray tracing in ray-centred coordinates	23
2.8.4	Second-order traveltimes approximation	23
2.8.5	Ray perturbation theory	24
3	Velocity model building using analytic diffraction traveltimes operators	27
3.1	The common-reflection-surface stack	27
3.1.1	Kinematic wavefield attributes	29
3.1.2	Practical aspects	30
3.1.3	Event-consistent smoothing and uncertainty estimation for CRS stacking parameters	31
3.1.4	Second-order diffraction traveltimes in terms of CRS stacking parameters	35
3.2	NIP-wave tomography	35
3.2.1	NIP-waves and velocities	35
3.2.2	The inverse problem	36
3.2.3	Data and model components	37
3.2.4	Data-driven determination of covariance operators	38
3.2.5	Automated picking strategy	39
3.2.6	The tomographic inversion algorithm	40
3.3	A velocity model building work flow based on second-order diffraction traveltimes functions	44
3.4	Data example	46
3.4.1	Data description	46
3.4.2	CRS stack results	47
4	Velocity model building from exact diffraction traveltimes	53
4.1	Motivation	53
4.2	Diffraction traveltimes as moveout function	54
4.2.1	Analytic example	56
4.3	The tomographic inversion scheme	61

4.3.1	Data and model components	61
4.3.2	Picking of time and time dip residuals	63
4.3.3	The tomographic inversion algorithm	64
4.3.4	Forward modelling and calculation of Fréchet derivatives	65
4.3.5	The structure of the tomographic matrix	66
4.4	Relationship between time dip residuals in the angle domain and in the midpoint-offset domain	67
4.5	Data example	68
5	Solution of the inverse problem and additional constraints	77
5.1	The solution of the inverse problem and its regularisation	77
5.2	Additional constraints	80
5.2.1	A-priori velocity information	80
5.2.2	Minimum velocity variation along reflectors	80
6	Applications	83
6.1	2D synthetic data example	83
6.1.1	Data description	83
6.1.2	CRS stack results	83
6.1.3	NIP-wave tomography and update	85
6.1.4	Depth migration results	88
6.1.5	Conclusions	89
6.2	3D real data example	89
6.2.1	Data description	90
6.2.2	CRS stack	90
6.2.3	Tomography	95
6.2.4	Migration	99
6.2.5	Conclusions	101
7	Summary and outlook	105
A	Velocity model in terms of B-splines	109

B Calculation of Fréchet derivatives	111
B.1 Fréchet derivatives in global Cartesian coordinates	111
B.1.1 Fréchet derivatives of \mathbf{x}_m and \mathbf{p}	111
B.1.2 Fréchet derivatives of \mathbf{M}	112
B.1.3 Fréchet derivatives of T	114
B.2 Fréchet derivatives in ray-centred coordinates	114
B.2.1 Fréchet derivatives for 2D NIP-wave tomography	115
B.2.2 Fréchet derivatives for the inversion scheme using model-based diffraction traveltime functions	119
B.3 Fréchet derivatives for additional constraints	120
B.3.1 A priori velocity information	120
B.3.2 Minimum velocity variation along reflectors	121
List of Figures	123
References	127
Danksagung	131
Lebenslauf	133

Chapter 1

Introduction

Reflection seismics play an important role in the exploration for natural resources, especially for fossil hydrocarbons. Whenever it is necessary to image complex geological structures to gain an understanding of the subsurface, the reflection seismic method is one of the best suited techniques at hand. Seismic data acquisition together with the processing and imaging steps build the reflection seismic method. The building of a model of wave propagation velocities which is used to transform the data, acquired as time series, into a detailed structured image of the subsurface is of special importance in this chain.

In this thesis methods for the construction of such velocity models are presented. In this chapter I give a short overview of seismic data acquisition, processing, and imaging. The role of velocity models for seismic imaging is discussed.

1.1 Seismic data acquisition

The exploration for hydrocarbon resources mainly takes place in sedimentary rocks located in the upper few kilometres of the Earth's crust. Surface seismic data, which will be considered in the following, are acquired on land as well as offshore. A complete seismic data acquisition usually consists of a number of single seismic experiments. Each such experiment comprises a seismic source and an array of receivers with distances up to several kilometres to the source. The distance between source and receiver is called offset. The source emits energy which propagates in the form of seismic waves through the subsurface. After reflection at interfaces (abrupt changes of the elastic properties) parts of these waves reach the surface where they are recorded by the receivers in terms of particle motion (or some function thereof, land acquisition) or pressure change (offshore acquisition) over a certain time span. The time series recorded at each receiver in this basic experiment is one trace of the final seismic dataset. This basic seismic experiment is called a common-shot (CS) experiment, the corresponding traces form a CS record. For a complete survey this basic experiment is repeated several times with different source-receiver geometries yielding a redundant illumination of the same subsurface structure. Due to this redundancy the resulting seismic dataset is called a multi-coverage dataset.

The most general distribution of sources and receivers is areal. This means that they cover a certain range of the surface. Assuming a flat acquisition surface each source and receiver position is given

by a two-dimensional vector in a coordinate system attached to the surface. Including recording time the resulting seismic dataset has five dimensions. The seismic experiment covers a three-dimensional subset of the earth in this case of 3D data acquisition. In marine data acquisition a vessel usually tows a small number of streamers containing the recording instruments. These streamers cover only a limited range of the surface. Therefore, the resulting dataset consists of traces which usually all have source-receiver azimuths in a small, limited range.

When the sources and receivers are constrained to a line on the surface, one speaks of 2D data acquisition. Generally, also 2D acquisition covers a 3D subset of the subsurface. However, this fact is usually neglected during processing. For a straight acquisition line exists only one source-receiver azimuth. Therefore, a dataset acquired in such a 2D way is three-dimensional.

Beside the above described data acquisition schemes, which all take place at the surface, there exist numerous special cases like vertical seismic profiling (VSP) and ocean bottom seismics (OBS), where sources and/or receivers are located in boreholes or on the seafloor. However, they are of no interest in the context of this thesis and, therefore, are not described here.

1.2 Seismic data processing

The recorded seismic data contain signals caused by a lot of different wave types as well as coherent and incoherent noise. Usually, only a very limited number of these signals is of interest for reflection seismics. One aim of seismic data processing is to suppress undesirable signals and coherent as well as random noise. The actual sequence of processing steps is different for each seismic dataset. However, a standard processing sequence involves at least the following steps (Yilmaz, 1987): preprocessing and filtering, deconvolution and multiple attenuation.

Laterally contiguous events are typically observed in seismic reflection data. These events are associated with wavefronts emerging at the surface and are commonly called seismic reflection events. The time span between the initiation of the source and the recording of the reflection signals is the traveltime of the associated wavefront. The goal of reflection seismic data processing is to enhance the reflection events to be used for the final construction of a structural image of the subsurface by a migration process. For imaging purposes, mostly events are used which are related to waves reflected exactly one time in the subsurface without a change of their polarisation. These events are commonly called primary reflection events.

Usually, seismic data are sorted with respect to the midpoint and distance between individual sources and receivers. In a two-dimensional Cartesian coordinate system attached to the flat horizontal measurement surface the midpoint and half-offset between a source at $\mathbf{x}_S = (x_S, y_S)^T$ and a receiver at $\mathbf{x}_G = (x_G, y_G)^T$ can be expressed by two-dimensional vectors. In this thesis, I use the following notation:

$$\mathbf{x}_m = \frac{1}{2}(\mathbf{x}_S + \mathbf{x}_G) \quad \text{and} \quad \mathbf{h} = \frac{1}{2}(\mathbf{x}_S - \mathbf{x}_G) . \quad (1.1)$$

A sketch of the above described situation is found in Figure 1.1. The azimuthal direction of both vectors is measured counterclockwise starting at the positive x-axis. Without loss of generality, I assume that the x-axis coincides with the seismic acquisition line in the 2D case. Thus, all vectors reduce to scalar quantities:

$$x_m = \frac{1}{2}(x_S + x_G) \quad \text{and} \quad h = \frac{1}{2}(x_S - x_G) . \quad (1.2)$$

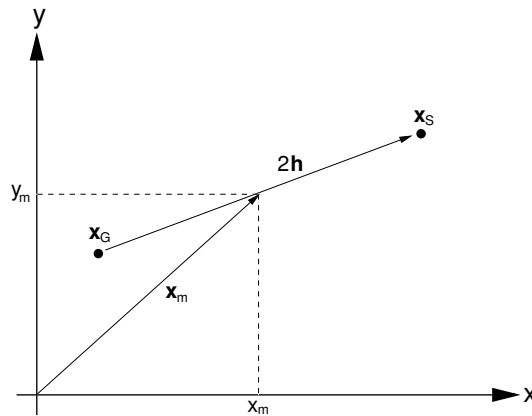


Figure 1.1: Definition of midpoint and half-offset coordinates.

All traces with the same midpoint between source and receiver coordinates form a common-midpoint (CMP) gather. In these gathers stacking velocity analysis is performed, i. e., a systematic analysis of the time-dependence of the coherent reflection events with respect to the source-receiver distance. The parameter describing this time dependence is called stacking velocity. The signals in the CMP gather are summed (stacked) after correction for the analysed time dependence to enhance coherent events, to increase the signal-to-noise ratio, and to reduce the amount of data to be handled in further processing steps. A simulated zero-offset (ZO) section (2D data acquisition) or volume (3D data acquisition) is obtained after stacking each CMP gather. All traces in this section/volume simulate recordings from an experiment with coincident source and receiver locations. The seismic data before stacking are commonly called prestack data.

Stacking velocity analysis explicitly makes use of the redundancy inherent in the multi-coverage dataset. Stacking velocity as a function of travelt ime and surface position can be used as basis for the construction of a subsurface wave propagation velocity model. This model is finally used in the depth migration step which transforms either the prestack data (prestack depth migration) or the simulated ZO section (poststack depth migration) into a structural image of the subsurface.

In classical stacking velocity analysis the time dependence of an event is only analysed with respect to the offset. This concept may be extended into the midpoint dimension as well which leads to stacking surfaces instead of stacking trajectories. One method of this type is the common-reflection-surface (CRS) stack technique. In this thesis it is used to extract information from the seismic data which is then used in a tomographic inversion scheme to determine a velocity model suitable for depth migration.

1.3 Seismic imaging and velocities

The imaging step is the core of the construction of a seismic subsurface image. The output of the imaging step may be either in the time or in the depth domain: the final image is either a function of surface coordinates and time or a function of surface coordinates and depth.

1.3.1 Time migration

A first image of the subsurface is obtained with the simulation of a ZO stack section by any kind of stacking procedure. This is not a depth image because the vertical axis is still time. Reflection events appear in a wrong lateral position and diffraction events are not yet focused.

A different time image is obtained using a process called time migration. This migration process ideally moves reflection events to their correct lateral position in the time domain and focuses diffraction events. However, the applicability of time migration strongly depends on the geological complexity of the area under investigation. Time migration may be applied either to a poststack section or the prestack data.

Both strategies, the simulation of a ZO section and time migration usually depend on a limited number of parameters (mostly just one), the best known ones being stacking and time migration velocity. These parameters can, in principle, be extracted from the seismic prestack data without knowledge of the true subsurface model although in their theoretical motivation some model assumptions are made like, for example, the assumption of a 1D model without lateral velocity variations. Both parameters bear a lot of information about the subsurface and are used in various kinds of inversion schemes for the construction of a subsurface wave propagation velocity model suitable for depth migration.

1.3.2 Depth imaging

A depth image is constructed from the seismic data by a process called depth migration. Various migration methods have been proposed: all are based on the wave equation and try to undo the wave propagation effects. This means that all these methods propagate or continue the recorded wavefield back into the subsurface. An overview of several migration methods can be found in [Yilmaz \(1987\)](#) and the references given therein. The final goal of depth migration is to achieve a depth image where all reflectors appear in their correct position and all diffracted energy is properly focused. To achieve this goal the distribution of seismic wave propagation velocities in the subsurface needs to be known. As this model is initially unknown it needs to be constructed from all information available: a priori geological knowledge about the target area, well logs, and the seismic reflection data themselves.

In principle, it is not possible to derive the exact wave propagation velocity model from the surface seismic data alone. However, using the redundancy inherent in a multi-coverage dataset, it is possible to find a model which is consistent with the acquired data. These models are commonly called macro velocity models. They are suitable for seismic imaging but are not suited for a detailed geological interpretation. Usually, a macro velocity model gets the better the more additional a priori information is available. In this thesis methods for the construction of such macro-velocity models are presented. If not stated otherwise the term ‘velocity model’ in the following refers to such macro-velocity models for depth migration.

1.4 Methods for velocity model estimation

Today, a wide variety of velocity model estimation techniques is available. Most of these methods are based on some criteria which describe the consistency between the model and the seismic data. They differ in the way these criteria are defined, measured, and used for the model determination.

Generally, the determination of a wave propagation velocity model is a very complicated process. Therefore, all methods make some assumptions about the subsurface in form of linearisations and approximations, or they are designed for a specific geological setting. As the kinematic properties of seismic wave propagation are nonlinear functions of the distribution of wave propagation velocity, virtually all methods work in an iterative way, either by updating a global model or by proceeding downwards layer by layer.

Due to the wide variety of methods available, it is hard to classify them. However, I would like to distinguish between two major groups of velocity model building techniques: methods that make use of the focusing properties of migration and methods using some form of traveltimes information extracted from the data. Many approaches combine aspects of both classes of methods. In the following, I will briefly discuss both cases. References given in this discussion are by no means complete. The reader interested in a comprehensive overview of velocity model building techniques is referred to [Robein \(2003\)](#).

1.4.1 Migration-based methods

Migration-based velocity model building techniques may be further subdivided into two classes: methods analysing the result of prestack depth migration as a function of offset or scattering angle and methods directly using the focusing properties of prestack migration.

The first class of methods is based on the following property: in a velocity model consistent with the data, seismic images obtained by prestack depth migration should be kinematically independent of source-receiver distance. This implies that reflection events have the same position and dip in all images obtained by depth migration of seismic data with different offsets. Any inconsistency of the migration velocity model may be observed as residual moveout in common-image gathers. A common-image gather (CIG) is the offset- or angle-dependent migration result for a fixed lateral image location. Generally, too low a migration velocity yields an upwards bending event with increasing offset while too high a velocity leads to the opposite behaviour. Only in the case of a consistent velocity model, a CIG shows flat events, i. e., events that are independent of offset. A consistent velocity model does not need to be identical with the true velocity distribution as the acquisition geometry restricts the maximum offset available. This means that the flatness of the events in a CIG can not be checked for all offsets for which specular reflections could have been recorded with an infinite extension of the acquisition geometry. The residual moveout in a CIG, i. e., the deviation of the events from the flat condition, may be measured and used to update the velocity model. For the case of simple subsurface models (no lateral velocity inhomogeneity, no reflector dips, and small source-receiver offsets), such updating techniques have been proposed by [Al-Yahya \(1989\)](#) and [Deregowski \(1990\)](#). These methods have been further developed by [Lafond and Levander \(1993\)](#) and [Liu and Bleistein \(1995\)](#) to include lateral velocity variations and dipping reflectors. The above mentioned methods make an analytic update to the velocity model and are usually applied in a layer-based fashion. There exist alternative techniques which transform the measured depth moveout into traveltimes residuals which are used in a tomographic update of the velocity model. These methods are described later.

The second class of migration-based methods is based on focusing analysis. These techniques use the fact that in a consistent velocity model all energy related to an image point collapses at zero offset and zero time using a downward continuation of shots and receivers. In case of an inconsistent velocity model, focusing occurs at a different depth and time. These deviations are used to calculate the true reflector depth and velocity value. However, this is strictly possible only for constant velocity media

and small reflector dips. This method has been introduced by [Jeannot and Faye \(1986\)](#), based on earlier work of [Yilmaz and Chambers \(1984\)](#). Attempts to overcome the above mentioned limitations due to simplifying assumptions have been made by [MacKay and Abma \(1992\)](#) and [Audebert and Diet \(1993\)](#).

Both classes of velocity model building techniques require the repeated application of prestack depth migration and, therefore, are computationally expensive. This holds especially for the 3D case.

1.4.2 Traveltime inversion

Other methods for velocity model building use some form of traveltime information extracted from the seismic data either in the time domain or in the depth domain after prestack depth migration. One of these approaches is reflection traveltime tomography: a model, built of reflector interfaces and velocities, is iteratively found by minimising the misfit between forward-modelled traveltimes and traveltimes of selected reflection events identified in the prestack data. Several traveltime inversion algorithms have been published in the past with early ones presented by [Sattlegger et al. \(1981\)](#) and [Gjøystdal and Ursin \(1981\)](#). Variations of classical reflection traveltime tomography, using various different model descriptions and parameterisations, are described in several publications, e. g., [Bishop et al. \(1985\)](#), [Farra and Madariaga \(1988\)](#), [Williamson \(1990\)](#), and [Stork and Clayton \(1991\)](#). All these methods are global methods. The associated nonlinear optimisation problem of minimising the misfit between forward-modelled and picked traveltimes is iteratively solved. In each iteration a global update of the velocity field and the interface structure is computed solving large, linearised systems of equations. Although generally yielding good results, traveltime tomography has one major drawback: the method requires a huge number of input data points (picks) for a successful inversion. The selection (picking) of the input traveltimes has to be performed in the prestack data along interpreted reflection events corresponding to the interfaces in the model description.

The identification and interpretation of reflection events is much easier in the post-migrated domain. Methods have been proposed which combine aspects of traveltime tomography and migration-based velocity model building techniques. Such techniques are described, e. g., in [Stork \(1992\)](#) and [Woodward et al. \(1998\)](#). Residual moveout measured in a CIG may be converted into traveltime residuals using the local reflector dip and velocity. These traveltime deviations are related to specular rays traced through the model and are used to calculate a global model update. Applied in this way, traveltime tomography is optimised for minimisation of the residual moveout in CIGs. Although the identification of reflection events is easier in the depth domain, the method is computationally very expensive as in each iteration a prestack depth migration is required.

All tomographic methods mentioned above only use traveltime residuals as input for the inversion scheme. There exist a number of methods which do not only use traveltime but also some of its derivatives. Probably the most famous of these methods is Dix inversion ([Dix, 1955](#)). Assuming a medium with horizontal reflectors and no lateral velocity variations (a 1D model) stacking velocity values obtained from stacking velocity analysis in a CMP gather may be transformed into a 1D interval velocity model. As will be described later, stacking velocity is a parameter characterising the curvature of a reflection event as a function of offset and, thus, an alternative description of the second derivative of traveltime with respect to source-receiver distance.

In addition to traveltimes, an approach called stereotomography introduced by [Billette and Lambaré \(1998\)](#) (see also [Billette et al., 2003](#)) uses their first partial derivatives (local time dip) at source and

receiver positions. Associated with each pick in the prestack data is a diffraction/reflection point in depth. The velocity distribution as well as all depth locations and local reflection/diffraction angles are updated simultaneously in an iterative way to minimise the misfit between picked and forward-modelled values of traveltimes, slopes at sources and receivers and the locations of sources and receivers. Each pick is treated independently of all others. Thus, there is no need to follow continuous coherent events in the prestack data, picking may be performed on locally coherent events, and no reflector interfaces need to be introduced in the model description. Meanwhile, there exist several extensions to and variations of this method. To simplify the picking procedure and to make the method more robust it has been formulated and applied in the poststack time domain by [Lavaud et al. \(2004\)](#). It has also been formulated as an updating technique ([Neckludov et al., 2006](#)). The input data for stereotomography may be picked from prestack depth migrated data as well. This method as well as the relationship between time and depth domain quantities is described in [Chauris et al. \(2002a,b\)](#).

Recently, a method using ZO traveltimes, their first derivatives with respect to midpoint coordinates, and their second derivatives with respect to offset has been introduced by [Duvencek and Hubral \(2002\)](#) and [Duvencek \(2003\)](#). It is thoroughly described in [Duvencek \(2004a\)](#) and [Duvencek \(2004b\)](#). The method has been introduced in the framework of seismic imaging based on the common-reflection-surface (CRS) stack method ([Jäger et al., 2001](#); [Hertweck et al., 2003](#)) and will be explained in detail later on as it plays a key role in this thesis. Picking of the input data points is performed in the poststack time domain which is easier and more robust than picking in the prestack data. There is no need to introduce reflector interfaces in the model description as each pick is treated independently of all others. Meanwhile, the method has found several applications (see, e. g., [Della Moretta et al., 2006](#)).

1.5 Migration velocity model building from diffraction traveltimes operators

Many of the velocity model building techniques mentioned in the previous section require either the repeated application of prestack migration or time consuming picking of traveltimes information in the seismic prestack data. A technique which avoids these drawbacks is the one described in [Duvencek \(2004a\)](#). It makes use of so-called kinematic wavefield attributes (see Section 3.1.1) which can be extracted from the seismic prestack data by means of the CRS stack method in a robust way. The kinematic wavefield attributes involved have a clear physical meaning. A subset of them gives a second-order approximation of the traveltimes due to a point source placed on a reflector in the subsurface. The complete theoretical background in the framework of CRS-based seismic imaging is comprehensively treated in [Duvencek \(2004a\)](#) and will be outlined later on.

In this thesis I make several extensions to this method: one focus is to further automatise the method, especially the extraction of input for the inversion from the CRS results. A second focus is the preconditioning of this input. Like in every tomographic inversion scheme which uses data of different physical dimensions the weighting of this data plays a key role. I present a method to determine these weights in a data-driven way from the output of the CRS stack method. Furthermore, the 3D inversion is extended to azimuth-dependent second-order traveltimes derivatives.

An analogy to the interpretation of a subset of the CRS stacking parameters in terms of the kinematic response of a point source is the interpretation as parameters describing second-order approximate

diffraction traveltimes functions. The method based on these approximate diffraction traveltimes operators yields accurate macro-velocity models in many cases. However, as the parameters only describe second-order approximations of the true diffraction traveltimes functions the obtained models are considered to be consistent with the prestack data up to second order, only. I introduce a new method for the 2D case which is based on model-dependent diffraction traveltimes functions and overcomes the limitation to second order. Only a small portion of the diffraction traveltimes operator calculated for a reflection point in depth is tangent to the reflection event in the prestack data. The trajectory in the prestack data volume containing all points of tangency, so-called stationary points, is called common-reflection-point (CRP) trajectory. The energy distributed along this trajectory is that energy ideally to be migrated to the corresponding image point in depth. To make sure that this energy is migrated to the same image point for all source-receiver offsets or scattering angles, respectively, a CRP gather must be flat like a CIG in conventional migration-based velocity analysis. The CRP gather is built of all traces in the prestack data containing a stationary point corresponding to the image point under consideration. However, a flat CRP gather does not imply a flat CIG in the case of sparsely selected analysis locations. I will introduce a criterion to measure the dip difference between the reflection event in the prestack data and the model-based diffraction traveltimes operator. For this purpose, the model-based diffraction traveltimes operator is considered as surface of recording time zero. This effectively results in a moveout correction of the prestack data. The final goal of the method is to ensure the kinematic equivalence between CRP gather and CIG for all considered image points. The method uses a subset of the input information required for the inversion scheme based on second-order diffraction traveltimes approximations and is implemented in a way that it can be used as a technique to update the models obtained with that method. Generally, the initial model for the new method needs to be much nearer to a consistent one than the initial model for the inversion scheme based on second-order approximations of diffraction traveltimes functions. For this reason, the new inversion scheme may be considered as a velocity model update technique.

1.6 Structure of the thesis

In the following chapters, the theory for the tomographic inversion based on second-order approximations of diffraction traveltimes functions is reviewed. A new method overcoming the limitation to second order is presented. Theoretical as well as practical aspects of both methods are discussed and their application to synthetic and real data examples is presented.

In this chapter, a short introduction to the basic principles of seismic imaging is given. A short description of seismic data acquisition and processing is followed by a summary of commonly applied techniques for migration velocity model building. Extensions to the tomographic inversion scheme presented by [Duvneck \(2004a\)](#) are described and a new velocity model building technique is proposed.

Chapter 2 covers the fundamentals of seismic wave propagation in isotropic media as far as necessary for an understanding of this thesis. An introduction to seismic ray theory is given and basic formulas involved in the implementation of the tomographic inversion schemes are derived. The developed theory forms the basis for relating data misfits to model changes in the inversion schemes.

Chapter 3 is concerned with the inversion scheme using second-order approximations of diffraction traveltimes functions. Many of the theoretical and practical issues of that method have already been discussed by [Duvneck \(2004a\)](#). I focus on several extensions to this method: an automated strategy

a	2-component vector
<u>a</u>	3-component vector
<u><u>a</u></u>	4-component vector
A	2×2 -matrix
<u>A</u>	3×3 -matrix
<u><u>A</u></u>	4×4 -matrix

Table 1.1: Notation for vectors and matrices in Cartesian coordinates. Vectors and matrices in other coordinate systems have special superscripts (see main text for details).

for input data determination within the framework of CRS-based seismic imaging will be presented together with a technique for the data-driven determination of input data uncertainties. These uncertainties may be used as covariance operators in the tomographic inversion algorithm. The 3D inversion scheme is extended to azimuth-dependent second-order traveltimes derivatives. A 2D synthetic data example highlighting the complete work flow for the velocity model building process is presented.

In **Chapter 4** I introduce a new tomographic inversion scheme based on model-dependent diffraction traveltimes functions for the 2D case. A method to measure the dip difference in the angle-domain between the reflection event in the prestack data and the forward-modelled diffraction traveltimes operator is described. Data and model components for the inversion are defined and the method is demonstrated with a synthetic data example.

The solution to the inverse problem to be solved in both presented inversion schemes is derived in **Chapter 5**. The formulation is general and valid for both inversion schemes. Additional constraints that may be applied are covered as well in this chapter.

Chapter 6 shows the results of applications of both inversion schemes. A 3D real data example for the inversion scheme based on analytic diffraction traveltimes functions is presented. Special emphasis is placed on a comparison between the models obtained with and without azimuth-dependent second-order traveltimes information. The complete work flow, starting with the CRS stack followed by both inversion schemes, is presented for a complex 2D synthetic data example. The data is too complex to be consistently described with second-order traveltimes approximations. Therefore, the inversion scheme using model-based diffraction traveltimes functions is applied as an updating technique. This leads to strong improvements in the final structural image.

Finally, **Chapter 7** summarises the presented work and gives an outlook on possible future research.

Aspects like the velocity model description in terms of B-splines and the calculation of Fréchet derivatives needed in the tomographic inversion schemes are treated in the appendices.

1.7 Notation

In various chapters of this thesis, matrix-vector calculus is used. If not stated otherwise, I use the notation summarised in Table 1.1. Vectors are set boldface with lowercase letters, whereas matrices are set with boldface uppercase letters. Vectors/matrices in other coordinates than Cartesian ones carry a special superscript enclosed in brackets: for example, a 2-component vector in ray-centred coordinates (still to be defined) is written as $\mathbf{a}^{(q)}$. An exception to this rule is the ray-centred coordinate vector \mathbf{q} defined in Section 2.8 which has no superscript.

Some equations are easier expressed in component notation. The indices are written as right-hand suffices with lowercase letters running from 1 to 3 and upper case letters taking the values 1 and 2. The Einstein summation convention is applied. Partial derivatives may also be indicated by subscripts, e. g, $a_{i,j}$ is a short notation for $\partial a_i / \partial x_j$. Partial time derivatives are indicated with a dot on top of the quantity, e. g., $\dot{\mathbf{a}}$ is a short notation for $\partial \mathbf{a} / \partial t$.

Matrices and vectors not following the above stated conventions are related to the solution of the inverse problem. They may have large dimensions depending on the number of picks and model parameters. They are set with uppercase and lowercase boldface letters, respectively.

Chapter 2

Ray and ray perturbation theory

The propagation of seismic waves is a considerably complex process governed by the elastodynamic wave equation. Using a high-frequency assumption, asymptotic solutions of the elastodynamic wave equation may be found. These are the eikonal and transport equations. They describe the kinematic and dynamic aspects of the wavefield. Whereas the transport equation is of no further interest in this thesis, the eikonal equation will be solved yielding the so-called ray tracing system.

In this thesis, paraxial ray theory and ray perturbation theory are of particular importance. They allow to examine the first-order effect of perturbations of model parameters (velocity, etc.) on the kinematic properties of wavefronts propagating along given ray trajectories. The theory will be developed as far as necessary for the formulation of the tomographic inversion schemes presented in later chapters.

Ray and ray perturbation theory are firstly introduced using a three-dimensional Cartesian coordinate system. This system is used for all 3D computations in this thesis. The results are then transformed to ray-centred coordinates (e. g., [Popov and Psencik, 1978](#); [Červený, 2001](#)) and reduced to the two-dimensional case. Most 2D computations will be performed in ray-centred coordinates.

Throughout this chapter, a smoothly varying, inhomogeneous, isotropic, elastic medium without discontinuities of the elastic properties is assumed. Certain applications of ray perturbation theory require continuous third spatial derivatives of the velocity field.

Further details on seismic wave propagation and ray theory may be found in many books, for example in [Aki and Richards \(1980\)](#), [Kravtsov and Orlov \(1990\)](#), and [Červený \(2001\)](#).

2.1 The elastodynamic wave equation

A common assumption in seismology is that the earth behaves as a linear elastic medium for small particle displacements. Seismic wave propagation is then described by the elastodynamic wave equation. The basic concepts and equations of linear elastodynamics as well as a derivation of the elastodynamic wave equation can be found in many books (see, among many others, [Aki and Richards, 1980](#); [Červený, 2001](#)).

Let $\underline{\mathbf{x}} = (x_1, x_2, x_3)$ denote the position of a certain particle in Cartesian coordinates. The particle's displacement vector with respect to its position in the unperturbed medium at time t is given by $\underline{\mathbf{u}}(\underline{\mathbf{x}}, t)$.

The elastodynamic equation for an unbound anisotropic, inhomogeneous, perfectly elastic medium in the absence of body forces reads

$$(c_{ijkl} u_{k,l})_{,j} = \rho \ddot{u}_i, \quad i = 1, 2, 3, \quad (2.1)$$

where c_{ijkl} denote the components of the elastic tensor connecting stress and strain. Both, the elastic tensor and the density may be spatially variable: $c_{ijkl} = c_{ijkl}(\mathbf{x})$ and $\rho = \rho(\mathbf{x})$.

The elastic tensor has 21 independent components in the most general case. This number reduces to just two in the case of isotropic media. The elastodynamic equation for isotropic media then reads

$$(\lambda u_{j,j})_{,i} + \left[\mu(u_{i,j} + u_{j,i}) \right]_{,j} + f_i = \rho \ddot{u}_i, \quad (2.2)$$

with λ and μ being the two Lamé parameters.

For homogeneous, isotropic media, where the elastic tensor and the density are independent of position, the solution of (2.2) are two completely decoupled wave types: the longitudinally polarised P-wave and the transversally polarised S-wave. However, for inhomogeneous, isotropic media, there is no general separation of the wavefield into independently propagating wave types. Nevertheless, for smoothly inhomogeneous media, high-frequency elastic waves approximately separate into P- and S-waves. In order to find a high-frequency solution to (2.2), a transient time signal of the form

$$\underline{\mathbf{u}}(\mathbf{x}, t) = \underline{\mathbf{U}}(\mathbf{x}) F(t - T(\mathbf{x})) \quad (2.3)$$

is considered, where $F(t - T(\mathbf{x}))$ represents a high-frequency, analytical signal (Červený, 2001). Essential here is that the Fourier spectrum $F(\omega)$ effectively vanishes for low frequencies. This implies

$$|F''(t)| \gg |F'(t)| \gg |F(t)|. \quad (2.4)$$

Inserting (2.3) into (2.2) yields

$$N_i(\underline{\mathbf{U}}) \ddot{F} - M_i(\underline{\mathbf{U}}) \dot{F} + L_i(\underline{\mathbf{U}}) F = 0, \quad (2.5)$$

where

$$\begin{aligned} N_i(\underline{\mathbf{U}}) &= -\rho U_i + (\lambda + \mu) U_j T_{,i} T_{,j} + \mu U_i T_{,j} T_{,j}, \\ M_i(\underline{\mathbf{U}}) &= (\lambda + \mu) \left[U_{j,i} T_{,j} + U_{j,j} T_{,i} + U_j T_{,jj} \right] + \mu \left[2U_{i,j} T_{,j} + U_i T_{,jj} \right] \\ &\quad + \lambda_i U_j T_{,j} + \mu_{,j} (U_i T_{,j} + U_j T_{,i}), \\ L_i(\underline{\mathbf{U}}) &= (\lambda + \mu) U_{j,ij} + \mu U_{i,jj} + \lambda_i U_{j,j} + \mu_{,j} (U_{i,j} + U_{j,i}). \end{aligned} \quad (2.6)$$

As equation (2.5) must be satisfied for all times, each of the individual terms must vanish independently. Considering equation (2.4) the term containing $L_i(\underline{\mathbf{U}})$ may be neglected with respect to the other terms under the high-frequency assumption. Thus, the following two equations remain:

$$N_i(\underline{\mathbf{U}}) = 0, \quad (2.7)$$

$$M_i(\underline{\mathbf{U}}) = 0. \quad (2.8)$$

Introducing the slowness vector $\underline{\mathbf{p}}$ as the gradient of $T(\mathbf{x})$, i. e. $\underline{\mathbf{p}} = \nabla T$, equation (2.7) may be reformulated to

$$(\Gamma_{ij} - \delta_{ij}) U_j = 0 \quad \text{with} \quad \Gamma_{ij} = \left[\frac{\lambda + \mu}{\rho} p_i p_j + \frac{\mu}{\rho} p_k p_k \delta_{ij} \right] \quad (2.9)$$

and

$$\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}.$$

Equation (2.9) states an eigenvalue problem for the matrix $\underline{\Gamma}$. Non-trivial solutions only exist if

$$\det(\Gamma_{ij} - \delta_{ij}) = \left(\frac{\mu}{\rho} p_k p_k - 1\right) \left(\frac{\lambda + \mu}{\rho} p_k p_k - 1\right)^2 = 0. \quad (2.10)$$

This leads to the possible solutions

$$p_k p_k = \left(\frac{\mu}{\rho}\right)^{-1} = v_S^{-2}, \quad (2.11a)$$

$$p_k p_k = \left(\frac{\lambda + \mu}{\rho}\right)^{-1} = v_P^{-2}. \quad (2.11b)$$

As can be seen from equations (2.11) two of the three eigenvalues coincide. However, as there are two different eigenvalues, two different wave types may be distinguished which propagate independently in isotropic, smoothly inhomogeneous media in the high-frequency approximation: P-waves polarised longitudinal in the direction of the slowness vector $\underline{\mathbf{p}} = \nabla T$ and S-waves polarised transversal in the plane perpendicular to $\underline{\mathbf{p}}$. The polarisation direction follows from the eigenvectors corresponding to the eigenvalues (2.11). The eikonal equations (2.11) control the kinematic characteristics of seismic wave propagation in smoothly inhomogeneous, isotropic media. From equation (2.8) the transport equations for P- and S-waves, controlling their dynamic aspects (amplitudes), may be derived. As dynamic aspects of the wavefield are of no interest in this thesis, equation (2.8) will not be further investigated.

The eikonal equations for P- and S-waves have the same form. For this reason, I will not distinguish between both wave types in the following derivations and will use a general wave propagation velocity v . In the next section a solution of the eikonal equation will be derived leading to the concept of rays.

2.2 The solution of the eikonal equation

In the preceding section the eikonal equation of high-frequency seismic waves propagating in smoothly inhomogeneous, isotropic media has been derived. In Cartesian coordinates it reads

$$p_k p_k = \frac{1}{v^2}, \quad (2.12)$$

with the slowness vector $\underline{\mathbf{p}}$ defined as the gradient of traveltime $T(\underline{\mathbf{x}})$. Equation (2.12) states a non-linear first-order partial differential equation for traveltime T and may be reformulated to

$$H(\underline{\mathbf{x}}, \underline{\mathbf{p}}) = (p_k p_k)^{\frac{1}{2}} - 1/v = 0, \quad (2.13)$$

where H is the Hamiltonian. This differential equation is usually solved using the method of characteristics (see, e. g., [Bleistein, 1984](#)) leading to the following characteristic system:

$$\frac{dx_i}{ds} = \frac{\partial H}{\partial p_i} = v p_i, \quad (2.14a)$$

$$\frac{dp_i}{ds} = -\frac{\partial H}{\partial x_i} = -\frac{1}{v^2} \frac{\partial v}{\partial x_i}, \quad (2.14b)$$

with s denoting the arclength along the ray. The six coupled equations (2.14) describe trajectories in a six-dimensional phase space with coordinates $(x_1, x_2, x_3, p_1, p_2, p_3)$, the characteristics of system (2.13). A projection into the spatial domain yields the 3D space trajectories referred to as rays. System (2.14) is therefore referred to as ray tracing system. In the case of isotropic media as considered in this thesis, rays are orthogonal to the wavefronts. The travelttime along a ray trajectory is determined using

$$\frac{dT}{ds} = \frac{1}{v}. \quad (2.15)$$

2.3 The reduced ray tracing system

In general, the number of first-order partial differential equations in the ray tracing system may be reduced from six to four. One of the three equations for p_i may be replaced by the eikonal equation. The corresponding coordinate x_i then replaces arclength s as independent variable along the ray trajectory. This procedure leads to so-called reduced ray tracing systems. As long as mainly near-vertical wave propagation is considered, it is obvious to choose the vertical component x_3 as independent variable in the ray tracing system. Solving (2.13) for p_3 and using $p_3 + H^R = 0$, yields the reduced Hamiltonian H^R :

$$H^R(x_1, x_2, x_3, p_1, p_2) = -\sqrt{1/v^2 - p_1^2 - p_2^2}. \quad (2.16)$$

The corresponding ray tracing system reads (Červený, 2001)

$$\frac{dx_I}{dx_3} = \frac{\partial H^R}{\partial p_I} = \frac{p_I}{p_3}, \quad (2.17a)$$

$$\frac{dp_I}{dx_3} = -\frac{\partial H^R}{\partial x_I} = -\frac{1}{v^3 p_3} \frac{\partial v}{\partial x_I}. \quad (2.17b)$$

The equation for calculating travelttime T along the ray trajectory is found combining equation (2.15) and the relation between arclength s and x_3 given by

$$\Delta x_3 = \Delta s \cos \theta = \Delta s p_3 v \quad (2.18)$$

where θ denotes the angle between the ray and the vertical. The final expression for calculation of travelttime T then reads

$$\frac{dT}{dx_3} = \frac{1}{v^2 p_3}. \quad (2.19)$$

Important for the validity of system (2.17) is that the ray has no turning point with respect to the x_3 -coordinate.

2.4 Paraxial ray tracing

Once a ray, referred to as central ray from now on, has been determined one may ask how a ray in its vicinity behaves. One possibility is to calculate that neighbouring ray trajectory changing the initial conditions of the central ray and solving the ray tracing system again. However, there is a simpler solution which allows to calculate an arbitrary number of neighbouring rays once some additional

calculations along the central ray have been performed. These additional calculations are given by the paraxial ray tracing system.

To obtain the paraxial ray tracing system it is assumed that a neighbouring ray may be described by small perturbations of the central ray's phase space coordinates:

$$x_I(x_3) = x_I^{(0)}(x_3) + \Delta x_I(x_3), \quad p_I(x_3) = p_I^{(0)}(x_3) + \Delta p_I(x_3), \quad (2.20)$$

with $x^{(0)}$ and $p^{(0)}$ denoting the unperturbed central ray. Like the central ray itself, all paraxial rays, i. e., rays in the close vicinity of the central ray, must satisfy the eikonal equation. Assuming small perturbations, the reduced Hamiltonian with perturbed coordinates may be obtained expanding the reduced Hamiltonian into a Taylor series up to the first order:

$$\begin{aligned} H^R(x_1, x_2, x_3, p_1, p_2) &\approx H^R(x_1^{(0)}, x_2^{(0)}, x_3, p_1^{(0)}, p_2^{(0)}) \\ &+ \left. \frac{\partial H^R}{\partial x_I} \right|_{x_I^{(0)}} \Delta x_I + \left. \frac{\partial H^R}{\partial p_I} \right|_{p_I^{(0)}} \Delta p_I. \end{aligned} \quad (2.21)$$

All partial derivatives in (2.21) are calculated on the central ray using the unperturbed reduced Hamiltonian. Inserting equation (2.21) into the ray tracing system (2.17) yields the paraxial ray tracing system

$$\begin{aligned} \frac{d}{dx_3} \Delta x_I &= \sum_{J=1}^2 \left(\frac{\partial^2 H^R}{\partial p_I \partial x_J} \Delta x_J + \frac{\partial^2 H^R}{\partial p_I \partial p_J} \Delta p_J \right), \\ \frac{d}{dx_3} \Delta p_I &= - \sum_{J=1}^2 \left(\frac{\partial^2 H^R}{\partial x_I \partial x_J} \Delta x_J + \frac{\partial^2 H^R}{\partial x_I \partial p_J} \Delta p_J \right). \end{aligned} \quad (2.22)$$

The linear system (2.22) may be formulated using matrix/vector notation:

$$\frac{d\tilde{\mathbf{w}}}{dx_3} = \tilde{\mathbf{S}} \Delta \tilde{\mathbf{w}}, \quad (2.23)$$

with the four component vector $\Delta \tilde{\mathbf{w}} = (\Delta x_1, \Delta x_2, \Delta p_1, \Delta p_2)^T$. The 4×4 matrix $\tilde{\mathbf{S}}$ is called the system matrix of the paraxial ray tracing system containing second derivatives of the reduced Hamiltonian H^R defined in equation (2.16).

2.4.1 The ray propagator matrix

The paraxial ray tracing system (2.23) is a linear system of four differential equations. Its general solution may be formulated in terms of a propagator matrix $\tilde{\mathbf{\Pi}}(x_3, x_{3_0})$ solving

$$\frac{d}{dx_3} \tilde{\mathbf{\Pi}} = \tilde{\mathbf{S}} \tilde{\mathbf{\Pi}} \quad (2.24)$$

and satisfying $\tilde{\mathbf{\Pi}}(x_{3_0}, x_{3_0}) = \tilde{\mathbf{I}}$ (the 4×4 identity matrix) at the starting point of the central ray. The solution of (2.23) for arbitrary initial conditions $\Delta \tilde{\mathbf{w}}$ can then be expressed as

$$\Delta \tilde{\mathbf{w}}(x_3) = \tilde{\mathbf{\Pi}}(x_3, x_{3_0}) \Delta \tilde{\mathbf{w}}(x_{3_0}). \quad (2.25)$$

The propagator matrix $\tilde{\mathbf{\Pi}}$ has a number of properties which are of special importance in the formulation of ray perturbation theory and for the implementation of the tomographic inversion schemes. They will be briefly discussed here. For a detailed treatment, the reader is referred to Červený (2001). Let me first write the propagator in a slightly different form introducing 2×2 sub-matrices:

$$\tilde{\mathbf{\Pi}}(x_3, x_{3_0}) = \begin{pmatrix} \mathbf{Q}_1(x_3, x_{3_0}) & \mathbf{Q}_2(x_3, x_{3_0}) \\ \mathbf{P}_1(x_3, x_{3_0}) & \mathbf{P}_2(x_3, x_{3_0}) \end{pmatrix}. \quad (2.26)$$

Symplecticity

The propagator matrix (2.26) is symplectic, i. e., it satisfies

$$\tilde{\mathbf{\Pi}}^T \tilde{\mathbf{J}} \tilde{\mathbf{\Pi}} = \tilde{\mathbf{J}} \quad \text{with} \quad \tilde{\mathbf{J}} = \begin{pmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{I} & \mathbf{0} \end{pmatrix}, \quad (2.27)$$

where \mathbf{I} is the 2×2 identity matrix and $\mathbf{0}$ is the 2×2 zero matrix, respectively.

Non-singularity

As a second important property, the propagator matrix satisfies

$$\det \tilde{\mathbf{\Pi}}(x_3, x_{3_0}) = 1. \quad (2.28)$$

Thus, the ray propagator matrix is non-singular along the entire central ray.

Chain rule

The third important property is the so-called chain rule. It states that in smooth media, if the second partial derivatives of the Hamiltonian are continuous functions of the independent variable, the propagator matrix for the complete central ray may be built up by multiplication of the propagator matrices of consecutive ray segments (see, e. g. Gilbert and Backus, 1966):

$$\begin{aligned} \tilde{\mathbf{\Pi}}(x_{3_n}, x_{3_0}) &= \tilde{\mathbf{\Pi}}(x_{3_n}, x_{3_{n-1}}) \tilde{\mathbf{\Pi}}(x_{3_{n-1}}, x_{3_{n-2}}) \dots \tilde{\mathbf{\Pi}}(x_{3_1}, x_{3_0}) \\ &= \prod_{i=n}^1 \tilde{\mathbf{\Pi}}(x_{3_i}, x_{3_{i-1}}), \end{aligned} \quad (2.29)$$

where the central ray has been split into n segments.

Inverse propagator matrix

Using the symplectic property and the chain rule the following relation for the inverse propagator matrix can be derived (Červený, 2001):

$$\tilde{\mathbf{\Pi}}^{-1}(x_3, x_{3_0}) = \begin{pmatrix} \mathbf{P}_2^T(x_3, x_{3_0}) & -\mathbf{Q}_2^T(x_3, x_{3_0}) \\ -\mathbf{P}_1^T(x_3, x_{3_0}) & \mathbf{Q}_1^T(x_3, x_{3_0}) \end{pmatrix}, \quad (2.30a)$$

$$\tilde{\mathbf{\Pi}}(x_{3_0}, x_3) = \tilde{\mathbf{\Pi}}^{-1}(x_3, x_{3_0}). \quad (2.30b)$$

All calculations are performed along the central ray from x_{3_0} to x_3 .

2.5 Dynamic ray tracing

A central ray is characterised by its initial conditions, i. e., by its initial position and slowness vector. All paraxial rays in its vicinity belonging to the same wavefront can be uniquely described by two parameters in the general 3D case. For the case of a point source these may, for example, be the two take-off angles in a spherical polar coordinate system fixed at the central ray's starting point. Dynamic ray tracing determines the first partial derivatives of the phase space coordinates of the central ray with respect to these parameters denoted by γ_1 and γ_2 in the following.

The partial derivatives of the phase space coordinates may be introduced as two 2×2 symmetric matrices:

$$Q_{IJ} = \frac{\partial x_I}{\partial \gamma_J}, \quad P_{IJ} = \frac{\partial p_I}{\partial \gamma_J}. \quad (2.31)$$

From equation (2.23) follows that these quantities satisfy the system of equations

$$\frac{d}{dx_3} \begin{pmatrix} \mathbf{Q} \\ \mathbf{P} \end{pmatrix} = \tilde{\mathbf{S}} \begin{pmatrix} \mathbf{Q} \\ \mathbf{P} \end{pmatrix} \quad (2.32)$$

defining the dynamic ray tracing system (Červený, 2001). This system is formally identical to the paraxial ray tracing system. However, the computed quantities have a different physical meaning. As the ray propagator matrix is a 4×4 identity matrix at the starting point of the central ray it can be immediately seen that the sub-matrix $(\mathbf{Q}_1, \mathbf{P}_1)^T$ of the ray propagator matrix is a solution of (2.32) for initial conditions $(\mathbf{I}, \mathbf{0})^T$, while $(\mathbf{Q}_2, \mathbf{P}_2)^T$ is a solution for initial conditions $(\mathbf{0}, \mathbf{I})^T$. The second initial condition is often called normalised point source condition (Červený, 2001).

2.6 Second-order traveltimes approximation

Let me assume that a central ray has been traced in a smoothly inhomogeneous, isotropic medium using the ray tracing system (2.17) and that the dynamic ray tracing system (2.32) has been solved along that ray. The 2×2 matrix \mathbf{M} of second derivatives of traveltimes with respect to coordinates x_1 and x_2 can then be determined using matrices \mathbf{Q} and \mathbf{P} defined in equation (2.31). Writing \mathbf{P} as

$$P_{IJ} = \frac{\partial p_I}{\partial \gamma_J} = \sum_{k=1}^2 \frac{\partial p_I}{\partial x_k} \frac{\partial x_k}{\partial \gamma_J} = \sum_{k=1}^2 \frac{\partial^2 t}{\partial x_I \partial x_k} \frac{\partial x_k}{\partial \gamma_J} = \sum_{k=1}^2 M_{Ik} Q_{kJ}, \quad (2.33)$$

the matrix \mathbf{M} is defined by

$$\mathbf{M} = \mathbf{P} \mathbf{Q}^{-1}. \quad (2.34)$$

Using this matrix, approximate second-order traveltimes of a specified wave can be calculated at arbitrary points near a central ray by dynamic ray tracing along that ray. The second-order traveltimes approximation for points $x_I + \Delta x_I$ in a horizontal plane with constant x_3 reads

$$t(x_1 + \Delta x_1, x_2 + \Delta x_2, x_3) = t(x_1, x_2, x_3) + \sum_{I=1}^2 p_I \Delta x_I + \sum_{I=1}^2 \sum_{J=1}^2 M_{IJ} \Delta x_I \Delta x_J. \quad (2.35)$$

Using the first partial derivatives of the velocity the above equation may be extended to include variations Δx_3 of the x_3 -coordinate (Červený, 2001). However, this extension is of no importance for this thesis.

2.7 Ray perturbation theory

The paraxial ray tracing system (2.23) describes ray trajectories in the close vicinity of the central ray. However, it might also be interpreted in a different way: it may be regarded as describing the first-order effect of perturbations $\Delta\tilde{\mathbf{w}}$ on the central ray trajectory. Thus, paraxial ray theory can be considered as a special case of ray perturbation theory. In addition to perturbations of the initial phase space coordinates, one might also consider the effect of perturbations of the Hamiltonian due to perturbations of the velocity $v(\mathbf{x})$ in the medium on the central ray trajectory. Furthermore, one may consider first-order effects on the propagator matrix along a central ray due to changes of the initial phase space coordinates as well as of the medium velocity. The resulting propagator matrix can be interpreted as the propagator matrix along a perturbed central ray trajectory. In this section, the expressions necessary for the tomographic inversion schemes presented later are derived. The development of ray perturbation theory given here follows the approach of [Farra and Madariaga \(1987\)](#).

Let me first introduce the perturbed Hamiltonian using a smoothly varying perturbation $\Delta v(\mathbf{x})$ of the background velocity:

$$v(\mathbf{x}) = v_0(\mathbf{x}) + \Delta v(\mathbf{x}) . \quad (2.36)$$

The corresponding perturbed Hamiltonian is then written as

$$H^R = H_0^R + \Delta H , \quad (2.37)$$

where H_0^R denotes the unperturbed Hamiltonian and

$$\Delta H^R = \frac{\partial H_0^R}{\partial v} \Delta v = \frac{1}{v^3 p_3} \Delta v . \quad (2.38)$$

Perturbations of the phase space coordinates may also be considered in (2.37). For that purpose, the perturbed phase space coordinates (2.20) are inserted into (2.37). Keeping only first-order terms in the perturbations, this yields

$$H^R = H_0^R + \frac{\partial H_0^R}{\partial x_I} \Delta x_I + \frac{\partial H_0^R}{\partial p_I} \Delta p_I + \frac{\partial H_0^R}{\partial v} \Delta v , \quad (2.39)$$

with all partial derivatives evaluated on the unperturbed central ray trajectory.

Let me now consider first-order effects on the central ray trajectory due to perturbations of its initial phase space coordinates and perturbations of the velocity. Therefore, I insert the perturbed Hamiltonian (2.39) into the right-hand side of ray tracing system (2.17) and keep only linear terms in the perturbations $\Delta x_1, \Delta x_2, \Delta p_1, \Delta p_2$, and Δv . This finally yields

$$\frac{d}{dx_3} \Delta\tilde{\mathbf{w}} = \tilde{\mathbf{S}} \Delta\tilde{\mathbf{w}} + \Delta\tilde{\mathbf{n}} , \quad (2.40)$$

with

$$\Delta\tilde{\mathbf{n}} = \left(\frac{\partial \Delta H^R}{\partial x_1}, \frac{\partial \Delta H^R}{\partial x_2}, -\frac{\partial \Delta H^R}{\partial p_1}, -\frac{\partial \Delta H^R}{\partial p_2} \right)^T . \quad (2.41)$$

This system of equations can again be solved in terms of a propagator ([Gilbert and Backus, 1966](#)). Using the propagator matrix $\tilde{\mathbf{\Pi}}$ the solution reads

$$\Delta\tilde{\mathbf{w}}(x_3) = \tilde{\mathbf{\Pi}}(x_3, x_{30}) \Delta\tilde{\mathbf{w}}(x_{30}) + \int_{x_{30}}^{x_3} \tilde{\mathbf{\Pi}}^{(x)}(x_3, x'_3) \Delta\tilde{\mathbf{n}}(x'_3) dx'_3 . \quad (2.42)$$

During ray tracing the propagator matrix from an intermediate position x'_3 to x_3 in the integral expression is usually unknown. Using the chain rule (2.29) and the inverse ray propagator (2.30) the propagator matrix in the integral may be entirely expressed with quantities known either during or after tracing the central ray:

$$\begin{aligned}\tilde{\mathbf{P}}^{(x)}(x_3, x'_3) &= \tilde{\mathbf{P}}^{(x)}(x_3, x_{3_0})\tilde{\mathbf{P}}^{(x)}(x_{3_0}, x'_3) \\ &= \tilde{\mathbf{P}}^{(x)}(x_3, x_{3_0})\tilde{\mathbf{P}}^{(x)-1}(x'_3, x_{3_0}).\end{aligned}\quad (2.43)$$

In general, the traveltime along the perturbed ray trajectory will differ from the one along the unperturbed central ray. Considering arclength s as independent variable along the ray trajectory, traveltime T along the unperturbed ray is calculated according to

$$T(s) = \int_{s_0}^s \frac{1}{v} ds'. \quad (2.44)$$

Thus, the perturbation of the traveltime due to changes in the velocity model is expressed as

$$\Delta T(s) = \int_{s_0}^s \Delta \left(\frac{1}{v} \right) ds' = - \int_{s_0}^s \frac{\Delta v}{v^2} ds'. \quad (2.45)$$

Using the relationship between arclength s and vertical coordinate x_3 ,

$$\Delta x_3 = \Delta s \cos(\alpha) = \Delta s p_3 v, \quad (2.46)$$

equation (2.45) transforms to the final expression in Cartesian coordinates for the calculation of traveltime changes due to changes in the velocity model:

$$\Delta T(x_3) = \int_{x_{3_0}}^{x_3} \frac{\Delta v}{v^3 p_3} dx'_3. \quad (2.47)$$

All calculations have to be performed along the unperturbed ray trajectory.

First-order effects on the traveltime due to a change of the starting point of the ray trajectory are given by

$$\Delta T(\Delta \mathbf{x}_0) = -\mathbf{p}_0 \cdot \Delta \mathbf{x}_0. \quad (2.48)$$

Due to Fermat's principle, perturbations of the initial slowness vector have no first-order effect on the traveltime.

Finally, the changes of the ray propagator matrix due to perturbations of the phase space coordinates and of the velocity are derived. For that purpose, the first order effects of these perturbations on the system matrix $\tilde{\mathbf{S}}$ need to be considered. The perturbed matrix $\tilde{\mathbf{S}}$ is expressed as

$$\tilde{\mathbf{S}} = \tilde{\mathbf{S}}_0 + \Delta \tilde{\mathbf{S}}. \quad (2.49)$$

In the following derivations, I distinguish between effects due to perturbations of the initial phase space coordinates and effects caused by perturbations of the velocity. Up to the first order in the perturbations, $\Delta \tilde{\mathbf{S}}$ can be decomposed yielding

$$\Delta \tilde{\mathbf{S}} = \Delta \tilde{\mathbf{S}}_1(\Delta \mathbf{w}) + \Delta \tilde{\mathbf{S}}_2(\Delta v). \quad (2.50)$$

Inserting the perturbed Hamiltonian (2.39) into the definition of $\tilde{\mathbf{S}}$ yields

$$\Delta\tilde{\mathbf{S}}_1(\Delta\tilde{\mathbf{w}}) = \sum_{l=1}^2 \left[\Delta x_l \frac{\partial}{\partial x_l} + \Delta p_l \frac{\partial}{\partial p_l} \right] \begin{pmatrix} \frac{\partial^2 H_0^R}{\partial p_1 \partial x_1} & \frac{\partial^2 H_0^R}{\partial p_1 \partial x_2} & \frac{\partial^2 H_0^R}{\partial p_1 \partial p_1} & \frac{\partial^2 H_0^R}{\partial p_1 \partial p_2} \\ \frac{\partial^2 H_0^R}{\partial p_2 \partial x_1} & \frac{\partial^2 H_0^R}{\partial p_2 \partial x_2} & \frac{\partial^2 H_0^R}{\partial p_2 \partial p_1} & \frac{\partial^2 H_0^R}{\partial p_2 \partial p_2} \\ -\frac{\partial^2 H_0^R}{\partial x_1 \partial x_1} & -\frac{\partial^2 H_0^R}{\partial x_1 \partial x_2} & -\frac{\partial^2 H_0^R}{\partial x_1 \partial p_1} & -\frac{\partial^2 H_0^R}{\partial x_1 \partial p_2} \\ -\frac{\partial^2 H_0^R}{\partial x_2 \partial x_1} & -\frac{\partial^2 H_0^R}{\partial x_2 \partial x_2} & -\frac{\partial^2 H_0^R}{\partial x_2 \partial p_1} & -\frac{\partial^2 H_0^R}{\partial x_2 \partial p_2} \end{pmatrix} \quad (2.51)$$

and

$$\Delta\tilde{\mathbf{S}}_2 = \begin{pmatrix} \frac{\partial^2 H_0^R}{\partial p_1 \partial x_1} & \frac{\partial^2 \Delta H_0^R}{\partial p_1 \partial x_2} & \frac{\partial^2 \Delta H_0^R}{\partial p_1 \partial p_1} & \frac{\partial^2 \Delta H_0^R}{\partial p_1 \partial p_2} \\ \frac{\partial^2 \Delta H_0^R}{\partial p_2 \partial x_1} & \frac{\partial^2 \Delta H_0^R}{\partial p_2 \partial x_2} & \frac{\partial^2 \Delta H_0^R}{\partial p_2 \partial p_1} & \frac{\partial^2 \Delta H_0^R}{\partial p_2 \partial p_2} \\ -\frac{\partial^2 \Delta H_0^R}{\partial x_1 \partial x_1} & -\frac{\partial^2 \Delta H_0^R}{\partial x_1 \partial x_2} & -\frac{\partial^2 \Delta H_0^R}{\partial x_1 \partial p_1} & -\frac{\partial^2 \Delta H_0^R}{\partial x_1 \partial p_2} \\ -\frac{\partial^2 \Delta H_0^R}{\partial x_2 \partial x_1} & -\frac{\partial^2 \Delta H_0^R}{\partial x_2 \partial x_2} & -\frac{\partial^2 \Delta H_0^R}{\partial x_2 \partial p_1} & -\frac{\partial^2 \Delta H_0^R}{\partial x_2 \partial p_2} \end{pmatrix}. \quad (2.52)$$

The matrix $\tilde{\mathbf{S}} = \tilde{\mathbf{S}}_0 + \Delta\tilde{\mathbf{S}}$ is used to describe paraxial rays in the close vicinity of the perturbed central ray:

$$\frac{d\Delta\tilde{\mathbf{w}}}{dx_3} = (\tilde{\mathbf{S}}_0 + \Delta\tilde{\mathbf{S}})\Delta\tilde{\mathbf{w}}, \quad (2.53)$$

where $\Delta\tilde{\mathbf{w}}$ denotes the perturbations of the phase space coordinates of the perturbed reference ray. Again, this system can be solved using a propagator matrix. This leads to

$$\Delta\tilde{\mathbf{w}}(x_3) = \tilde{\mathbf{\Pi}}_0^{(x)}(x_3, x_{3_0})\Delta\tilde{\mathbf{w}}(x_{3_0}) + \int_{x_{3_0}}^{x_3} \tilde{\mathbf{\Pi}}_0(x_3, x'_3)\Delta\tilde{\mathbf{S}}(x'_3)\Delta\tilde{\mathbf{w}}(x'_3) dx'_3. \quad (2.54)$$

Up to first order in the perturbation $\Delta\tilde{\mathbf{S}}$ equation (2.54) yields (Farra and Madariaga, 1987)

$$\Delta\tilde{\mathbf{w}}(x_3) = \left[\tilde{\mathbf{\Pi}}_0(x_3, x_{3_0}) + \Delta\tilde{\mathbf{\Pi}}(x_3, x_{3_0}) \right] \Delta\tilde{\mathbf{w}}(x_{3_0}), \quad (2.55)$$

where the perturbation of the ray propagator matrix is given by

$$\Delta\tilde{\mathbf{\Pi}}(x_3, x_{3_0}) = \int_{x_{3_0}}^{x_3} \tilde{\mathbf{\Pi}}_0(x_3, x'_3)\Delta\tilde{\mathbf{S}}(x'_3)\tilde{\mathbf{\Pi}}_0(x'_3, x_{3_0}) dx'_3. \quad (2.56)$$

For the actual computation of equation (2.56) expression (2.43) is substituted for $\tilde{\mathbf{\Pi}}_0(x'_3, x_{3_0})$ in the above integral expression.

All derivations up to now are based on the Hamiltonian approach using the reduced Hamiltonian defined in equation (2.16). The derived equations will be used for the 3D tomographic inversion schemes. However, they could as well be derived in any arbitrary curvilinear coordinate system. Especially useful, and later on used for 2D computations, is the ray-centred coordinate system.

2.8 Ray-centred coordinates

The theory presented above in Cartesian coordinates is translated to ray-centred coordinates in this section. The ray-centred coordinate system is introduced for the general 3D case. As this coordinate

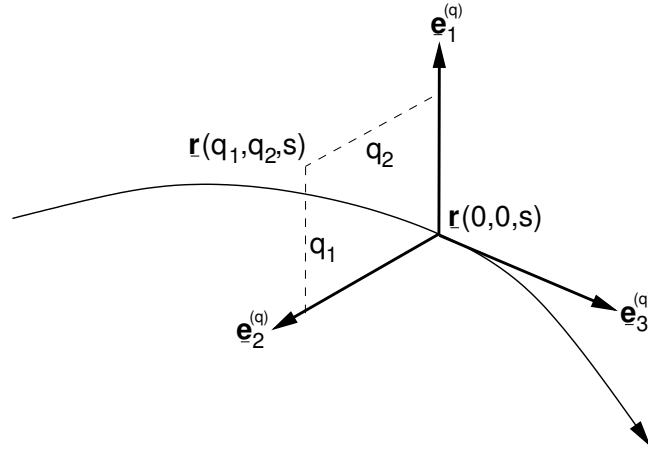


Figure 2.1: Ray-centred coordinates of a point in the vicinity of a central ray trajectory. See main text for details.

system is only used for the 2D inversion schemes in the scope of this thesis, all further results are reduced to the 2D case.

The ray-centred coordinate system $\underline{\mathbf{q}} = (q_1, q_2, q_3)$ is directly connected with a ray: one coordinate, here I choose q_3 , is the monotonic parameter along the ray. Thus, the ray itself is one of the coordinate system axes. Without loss of generality, I choose arclength s to be the monotonic parameter along the ray, i. e., $q_3 = s$. The unit basis vector $\underline{\mathbf{e}}_3^{(q)}$ is a vector tangent to the ray in each point along the ray trajectory. The remaining two basis vectors are chosen to build a right-handed orthogonal triplet with $\underline{\mathbf{e}}_3^{(q)}$ and are, thus, situated in a plane perpendicular to the ray. Note that all basis vectors are functions of arclength s and have to be computed along the ray according to the following system of equations (Červený, 2001):

$$\frac{\underline{\mathbf{e}}_I^{(q)}}{ds} = (\underline{\mathbf{e}}_I^{(q)} \nabla_V) \underline{\mathbf{p}}, \quad I = 1, 2. \quad (2.57)$$

Once the basis vectors form a mutually orthogonal triplet at the initial point of the ray they keep this characteristic along the whole ray trajectory. A point R' outside the ray trajectory is described in ray-centred coordinates as follows:

$$\underline{\mathbf{r}}(q_1, q_2, s) = \underline{\mathbf{r}}(0, 0, s) + q_1 \underline{\mathbf{e}}_1^{(q)}(s) + q_2 \underline{\mathbf{e}}_2^{(q)}(s). \quad (2.58)$$

Keeping arclength s fixed, only points in the plane perpendicular to the ray can be described. In order to describe points outside of this plane in ray-centred coordinates one has to move along the ray to change the origin of the ray-centred coordinate system. A sketch of the situation described here is given in Figure 2.1. The coordinates $\underline{\mathbf{r}}(q_1, q_2, s)$ of a point in the vicinity of a ray are uniquely defined if there is only one plane perpendicular to the ray which contains that point. If several such planes can be constructed the point is no longer uniquely defined. The validity of the ray-centred coordinate system is restricted to that range around the ray, in which points may be uniquely described. Obviously, this range of validity depends on the curvature of the ray and the model complexity.

The infinitesimal line segment dl in ray-centred coordinates satisfies (Červený, 2001)

$$(dl)^2 = d\underline{\mathbf{r}} \cdot d\underline{\mathbf{r}} = dq_1^2 + dq_2^2 + h^2 ds^2 \quad (2.59)$$

with

$$h = 1 + \frac{1}{v} \frac{\partial v}{\partial q_I} \Big|_{q_1=q_2=0} q_I . \quad (2.60)$$

2.8.1 Transformation to Cartesian coordinates

As the metric tensor of the ray-centred coordinate system contains the model dependent term (2.60), the transformation from ray-centred to global Cartesian coordinates is generally complex. It may be considerably simplified introducing a local ray-centred Cartesian coordinate system with its origin at s on the central ray. Let the local ray-centred Cartesian coordinates be denoted by (y_1, y_2, y_3) . The directions of the y_1 -axis and y_2 -axis coincide with the basis vectors $\underline{\mathbf{e}}_1^{(q)}$ and $\underline{\mathbf{e}}_2^{(q)}$ of the ray-centred coordinate system and the y_3 -axis is tangent to ray at s . When the origin of the local ray-centred Cartesian coordinate system is located at $\underline{\mathbf{x}}_0$ in the global Cartesian coordinate system, the coordinates of a point in local ray-centred Cartesian coordinates are transformed to the global Cartesian coordinate system by

$$(\underline{\mathbf{x}} - \underline{\mathbf{x}}_0) = \underline{\mathbf{H}} \underline{\mathbf{y}} , \quad (2.61)$$

with

$$H_{kl}(s) = \underline{\mathbf{e}}_k \cdot \underline{\mathbf{e}}_l^{(q)}(s) = \frac{\partial x_k}{\partial q_l} \Big|_s = \frac{\partial q_l}{\partial x_k} \Big|_s . \quad (2.62)$$

Obviously, the columns of matrix $\underline{\mathbf{H}}$ are given by the Cartesian components of the basis vectors of the ray-centred coordinate system at the considered point on the central ray trajectory.

2.8.2 Paraxial ray tracing in ray-centred coordinates

In the 2D case, wave propagation is confined to a plane which I choose, without loss of generality, to be the plane defined by $x_2 = 0$. Thus, ray trajectories may be described by two spatial coordinates. Consequently, only two ray-centred coordinates remain: arclength s measured along the ray and q measured vertical to the ray in the wave propagation plane.

The eikonal equation in ray-centred coordinates for the 2D case reads (see, e. g., Červený, 2001)

$$\left(\frac{\partial T}{\partial q} \right)^2 + \frac{1}{h^2} \left(\frac{\partial T}{\partial s} \right)^2 = \frac{1}{v^2(q, s)} . \quad (2.63)$$

Introducing the slowness component $p^{(q)}$ with

$$p^{(q)} = \frac{\partial T}{\partial q} , \quad (2.64)$$

the reduced Hamiltonian may be written as

$$H^R = -h \sqrt{\frac{1}{v^2} - p^{(q)2}} . \quad (2.65)$$

From (2.65) the paraxial ray tracing system in ray-centred coordinates is derived. As the central ray is described by $q = 0$, the perturbation of the ray trajectory is directly given by q . The paraxial ray tracing system in ray-centred coordinates finally reads (Červený, 2001)

$$\frac{dq}{ds} = v(0, s) p^{(q)} , \quad \frac{dp^{(q)}}{ds} = -\frac{1}{v^2(0, s)} \frac{\partial^2 v(0, s)}{\partial q \partial q} q . \quad (2.66)$$

The velocity and all its derivatives have to be calculated on the central ray. Introducing the vector $\mathbf{w}^{(q)} = (q, p^{(q)})$ the paraxial ray tracing system may be formulated in matrix/vector notation:

$$\frac{d\mathbf{w}^{(q)}}{ds} = \mathbf{S}^{(q)}\mathbf{w}^{(q)}, \quad (2.67)$$

where $\mathbf{S}^{(q)}$ represents the 2×2 system matrix. It reads

$$\mathbf{S}^{(q)} = \begin{pmatrix} 0 & v(0, s) \\ -\frac{1}{v^2(0, s)} \frac{\partial^2 v(0, s)}{\partial q^2} & 0 \end{pmatrix}. \quad (2.68)$$

The 2×2 ray propagator matrix associated with system (2.67) is denoted by

$$\mathbf{\Pi}^{(q)}(s, s_0) = \begin{pmatrix} Q_1^{(q)}(s, s_0) & Q_2^{(q)}(s, s_0) \\ P_1^{(q)}(s, s_0) & P_2^{(q)}(s, s_0) \end{pmatrix}. \quad (2.69)$$

2.8.3 Dynamic ray tracing in ray-centred coordinates

Again, the dynamic ray tracing system is found by introducing the first partial derivatives of phase space coordinates q and $p^{(q)}$ with respect to an initial parameter γ of the central ray, keeping other initial parameters fixed:

$$Q^{(q)} = \frac{\partial q}{\partial \gamma}, \quad P^{(q)} = \frac{\partial p^{(q)}}{\partial \gamma}. \quad (2.70)$$

These quantities satisfy the dynamic ray tracing system

$$\frac{d}{ds} \begin{pmatrix} Q^{(q)} \\ P^{(q)} \end{pmatrix} = \mathbf{S}^{(q)} \begin{pmatrix} Q^{(q)} \\ P^{(q)} \end{pmatrix}. \quad (2.71)$$

The system matrix $\mathbf{S}^{(q)}$ has been defined in equation (2.68). Associated with this system is the propagator matrix $\mathbf{\Pi}^{(q)}(s, s_0)$ already defined in equation (2.69).

In analogy to the 3D case in Cartesian coordinates, $(Q_1^{(q)}, P_1^{(q)})^T$ is a solution for initial conditions $(1, 0)^T$, known as initial normalised plane-wave conditions in the case of ray-centred coordinates (Červený, 2001), and $(Q_2^{(q)}, P_2^{(q)})^T$ is a solution for initial normalised point source conditions $(0, 1)^T$.

2.8.4 Second-order traveltimes approximation

Once a central ray trajectory has been found and the dynamic ray-tracing system (2.71) has been solved along it, approximate second-order traveltimes of points on a specific wavefront in the vicinity of the central ray are given by

$$t(q, s) = t(0, s) + M^{(q)} q^2. \quad (2.72)$$

The second derivative of traveltimes with respect to q for any specified wave is denoted with $M^{(q)}$:

$$\frac{\partial^2 t}{\partial q^2} = M^{(q)} = \frac{P}{Q}. \quad (2.73)$$

The case of a wave due a point source at $s = s_0$ is of special importance in this thesis. In that case,

$$M^{(q)} = \frac{P_2}{Q_2} . \quad (2.74)$$

Second-order approximations of traveltimes at arbitrary points in the close vicinity of the central ray are easily expressed in local ray-centred Cartesian coordinates. The traveltime approximation for the 2D case ($y_2 = 0$) reads

$$t(y_1, y_3) = t(0, 0) + \frac{1}{v(s)}y_3 + \frac{1}{2}(y_1, y_3) \begin{pmatrix} M^{(q)} & -\frac{1}{v^2} \frac{\partial v}{\partial y_1} \\ -\frac{1}{v^2} \frac{\partial v}{\partial y_1} & -\frac{1}{v^2} \frac{\partial v}{\partial y_3} \end{pmatrix} \begin{pmatrix} y_1 \\ y_3 \end{pmatrix} . \quad (2.75)$$

It needs the first partial derivatives of the velocity at the location on the central ray where the local ray-centred Cartesian coordinate system is fixed.

2.8.5 Ray perturbation theory

In analogy to the 3D case in Cartesian coordinates, the paraxial ray tracing system (2.66) can be considered as a special case of ray perturbation theory: it describes first-order changes $\mathbf{w}^{(q)}$ of the phase space coordinates of the central ray due to perturbations of its initial conditions. However, these are not the only effects one might consider. In this thesis, first-order effects of perturbations of the Hamiltonian (2.65) due to perturbations of the velocity as well as first-order effects on the ray propagator matrix due to perturbations of the velocity and the phase space coordinates are of special importance. Here, I will give the corresponding equations of ray perturbation theory, specialised to 2D ray centred coordinates. The perturbed velocity along the central ray trajectory is described by $v(0, s) = v_0(0, s) + \Delta v(0, s)$, where v_0 denotes the unperturbed velocity.

The perturbations of the ray-centred coordinate q and its associated slowness $p^{(q)}$ at the end point of the central ray due to perturbations of the initial phase space coordinates and the velocity model are solutions to the inhomogeneous dynamic ray tracing system

$$\mathbf{w}^{(q)} = \mathbf{S}^{(q)} \mathbf{w}^{(q)} + \Delta \mathbf{n}^{(q)} , \quad (2.76)$$

and are given by an expression analogous to equation (2.42):

$$\mathbf{w}^{(q)}(s) = \mathbf{\Pi}^{(q)}(s, s_0) \mathbf{w}^{(q)}(s_0) + \mathbf{\Pi}^{(q)}(s, s_0) \int_{s_0}^s \mathbf{\Pi}^{(q)-1}(s', s_0) \Delta \mathbf{n}^{(q)}(s', \Delta v(s')) ds' , \quad (2.77)$$

where the chain rule (2.29) has been applied. The vector $\Delta \mathbf{n}^{(q)}$ is given by

$$\Delta \mathbf{n}^{(q)} = \left(\frac{\partial \Delta H^R}{\partial p} , -\frac{\partial \Delta H^R}{\partial q} \right)^T , \quad (2.78)$$

with

$$\Delta H^R = \frac{\partial H^R}{\partial v} \Delta v .$$

Changes of the traveltime t due to perturbations of the velocity model are calculated according to integral (2.45). Together with perturbations of the ray's starting point parallel to the central ray trajectory, the final expression for traveltime perturbations reads

$$\Delta t(s) = - \int_{s_0}^s \frac{\Delta v}{v^2(s')} ds' - \frac{\sin \theta}{v(s_0)} \Delta x_0 - \frac{\cos \theta}{v(s_0)} \Delta z_0 . \quad (2.79)$$

The perturbations of the elements of the ray propagator matrix are computed along the central ray according to

$$\Delta \mathbf{\Pi}^{(q)}(s, s_0) = \mathbf{\Pi}^{(q)}(s, s_0) \int_{s_0}^s \mathbf{\Pi}^{(q)-1}(s', s_0) \Delta \mathbf{S}^{(q)}(s') \mathbf{\Pi}^{(q)}(s', s_0) ds' , \quad (2.80)$$

where the perturbation of system matrix $\mathbf{S}^{(q)}$ is, like in the 3D Cartesian case, divided in a part due to perturbations of the initial phase space coordinates of the central ray and a second part due to perturbations of the velocity model:

$$\Delta \mathbf{S}^{(q)}(s) = \Delta \mathbf{S}_1^{(q)}(\Delta v) + \Delta \mathbf{S}_2^{(q)}(\Delta q, \Delta p) .$$

I have used the notation Δq and Δp to distinguish between perturbations of the perturbed and unperturbed central ray.

Chapter 3

Velocity model building using analytic diffraction traveltimes operators

In this chapter, a method for velocity model building is described which makes use of second-order approximations of diffraction traveltimes operators. These approximate diffraction traveltimes operators are parametrised with first and second traveltimes derivatives which are determined from the seismic prestack data. The method presented here is based on the work of [Duveneck \(2004a\)](#) who developed the method in the framework of common-reflection-surface stack-based ([Müller, 1999](#)) seismic imaging. I will extend this method to azimuth-dependent second-order traveltimes derivatives in the 3D case.

A second focus of this chapter is the extraction as well as pre-conditioning of the input data for the inversion scheme. A new method for the data-driven determination of uncertainties of the input data within the framework of CRS-based seismic imaging is presented. These uncertainties are determined in a statistical manner and are applied as covariance operators in the inversion scheme.

3.1 The common-reflection-surface stack

Parameters for a second-order approximation of diffraction traveltimes operators may be extracted in several ways from the seismic prestack data. In this thesis I make use of the common-reflection-surface stack technique for this purpose. Alternatively, conventional stacking velocity analysis followed by a local slant stack technique in the ZO stacked section/volume could be performed.

The CRS stack (e. g., [Mann et al., 1999](#); [Jäger et al., 2001](#); [Mann, 2002](#)) is a method for the simulation of stacked ZO sections (2D case) or volumes (3D case) with an improved signal-to-noise (S/N) ratio. The redundancy inherent in the prestack multi-coverage data set (see [Chapter 1](#)) is used to extract several quantities describing the CRS stacking operator for each zero-offset output sample: these so-called kinematic wavefield attributes have a clear physical meaning and can be utilised to parametrise a second-order approximation of the diffraction traveltimes operator associated with each zero-offset sample.

The CRS operator itself gives a second-order approximation of traveltimes associated with reflections on a small element of a subsurface interface. This traveltimes approximation can be derived without

explicit knowledge about the subsurface structure and velocity model. For a comprehensive geometrical derivation of the CRS stack operator I refer to Höcht (2002). Unlike the well-known CMP stack operator (Mayne, 1962) which is confined to the time-offset domain the CRS operator includes also the midpoint dimension. Thus, the CRS operator describes an entire stacking surface.

Using the midpoint/half-offset coordinates \mathbf{x}_m and \mathbf{h} already introduced in Chapter 1 the operator for the 3D case reads

$$t^2(\mathbf{x}_m + \Delta\mathbf{x}_m, \mathbf{h}) = (t_0 + 2\mathbf{p}\Delta\mathbf{x}_m)^2 + 2t_0 \left(\Delta\mathbf{x}_m^T \mathbf{N} \Delta\mathbf{x}_m + \mathbf{h}^T \mathbf{M} \mathbf{h} \right). \quad (3.1)$$

It depends on eight stacking parameters: a two-component vector \mathbf{p} and two symmetric 2×2 matrices \mathbf{M} and \mathbf{N} . These stacking parameters are first and second traveltimes derivatives with respect to midpoint and half-offset coordinates. Assuming the near-surface velocity v_0 to be known they may be expressed in terms of the so-called kinematic wavefield attributes:

$$2\mathbf{p} = \left(\frac{\partial t}{\partial x_m}, \frac{\partial t}{\partial y_m} \right)^T = \frac{2}{v_0} (\cos\alpha \sin\beta, \sin\alpha \sin\beta)^T, \quad (3.2a)$$

$$\mathbf{N} = \begin{pmatrix} \frac{\partial^2 t}{\partial x_m^2} & \frac{\partial^2 t}{\partial x_m \partial y_m} \\ \frac{\partial^2 t}{\partial y_m \partial x_m} & \frac{\partial^2 t}{\partial y_m^2} \end{pmatrix} = \frac{1}{v_0} \mathbf{H} \mathbf{K}_N \mathbf{H}^T, \quad (3.2b)$$

$$\mathbf{M} = \begin{pmatrix} \frac{\partial^2 t}{\partial h_x^2} & \frac{\partial^2 t}{\partial h_x \partial h_y} \\ \frac{\partial^2 t}{\partial h_y \partial h_x} & \frac{\partial^2 t}{\partial h_y^2} \end{pmatrix} = \frac{1}{v_0} \mathbf{H} \mathbf{K}_{NIP} \mathbf{H}^T. \quad (3.2c)$$

The meaning of the kinematic wavefield attributes α , β , \mathbf{K}_{NIP} , and \mathbf{K}_N is explained in section 3.1.1. Their derivation is based on the concept of normal rays which connect the surface location \mathbf{x}_m with a subsurface reflection point. Matrix \mathbf{H} describes the transformation from local ray-centred Cartesian coordinates to the global Cartesian coordinate system at the normal ray emergence location \mathbf{x}_m (see Section 2.8.1).

In the 2D case, the CRS stacking operator only depends on three parameters. Expressed in terms of kinematic wavefield attributes it reads

$$t^2(x_m, h) = \left(t_0 + \frac{2 \sin\beta}{v_0} \Delta x_m \right)^2 + \frac{2t_0 \cos^2\beta}{v_0} (K_N \Delta x_m^2 + K_{NIP} h^2). \quad (3.3)$$

Like in the 3D case there is a relationship between traveltimes derivatives and wavefield attributes:

$$p = \frac{\sin\beta}{v_0}, \quad (3.4a)$$

$$N = \frac{\cos^2\beta}{v_0} K_N, \quad (3.4b)$$

$$M = \frac{\cos^2\beta}{v_0} K_{NIP}. \quad (3.4c)$$

All traveltimes derivatives and corresponding kinematic wavefield attributes are scalar quantities in the 2D case.

Without extension into the midpoint dimension, the CRS stacking operator reduces to the well known CMP stack formula (Mayne, 1962). For the 2D case it reads

$$t^2(h) = t_0^2 + \frac{4h^2}{v_{NMO}^2}, \quad (3.5)$$

where v_{NMO} is the normal-moveout (NMO) velocity. This equation gives a second-order approximation for traveltimes in a CMP ensemble valid for arbitrary models with smooth lateral velocity variations and smoothly curved interfaces. The relationship between NMO velocity and CRS stacking parameters for the general 3D case is

$$v_{\text{NMO}}^{-2}(\phi) = \frac{t_0}{2} \check{\mathbf{e}}_\phi^T \mathbf{M} \check{\mathbf{e}}_\phi, \quad (3.6)$$

where the two-component unit vector $\check{\mathbf{e}}_\phi = (\cos \phi, \sin \phi)^T$ describes the azimuthal direction in the measurement surface in which the second traveltime derivative is evaluated. In the case that the acquisition geometry allows only offset vectors restricted to a small azimuth range near a given direction ϕ , the component M_ϕ of matrix \mathbf{M} associated with direction ϕ is the only one which can be determined from the seismic data:

$$M_\phi = \check{\mathbf{e}}_\phi^T \mathbf{M} \check{\mathbf{e}}_\phi. \quad (3.7)$$

In the 2D case, the relationship between normal-moveout velocity and kinematic wavefield attributes reads

$$v_{\text{NMO}}^2 = \frac{2v_0}{t_0 K_{\text{NIP}} \cos^2 \beta}. \quad (3.8)$$

In practice, stacking velocities derived from the seismic data by stacking velocity analysis or the CRS method do not coincide with NMO velocities defined by the second-order traveltime approximation (3.5). Stacking velocities are best-fit parameters describing the moveout in CMP ensembles. Thus, they depend on the applied apertures in the search step. Normal-moveout velocities appear as second-order coefficients in the series expansion for traveltimes in a CMP ensemble. The deviation between both velocities is called spread-length bias (see, e. g., [Al-Chalabi, 1973](#); [Hubral and Krey, 1980](#)). Thus, normal-moveout velocities calculated from best fit CRS stacking parameters only have an approximate meaning.

The absence of a linear term in the offset dimension of above traveltime formulas is a consequence of the normal ray approach used in the explanation of the kinematic wavefield attributes. The normal ray hits the reflector in the subsurface normally. First-order terms may exist for non-normal ZO rays. As long as they depend only on the absolute value $|h|$ of the half-offset, this does not collide with the principle of reciprocity which states that from a kinematic point of view source and receiver may be interchanged. However, in this thesis reflection events related to non-normal ZO rays are not considered.

In contrast to homogeneous models, the reflection events in a CMP gather generally do not correspond to a common reflection point in depth for heterogeneous velocity models and/or dipping reflector interfaces. This phenomenon is called reflection point dispersal.

3.1.1 Kinematic wavefield attributes

The coefficients in the second-order traveltime approximation (3.1) can be related to kinematic properties of two hypothetical wavefronts emerging at the surface at \mathbf{x}_m . The stacking parameters \mathbf{p} and \mathbf{N} describe the first and second horizontal traveltime derivatives related to an emerging wavefront at \mathbf{x}_m which stems from an exploding reflector segment placed at the normal incidence point (NIP) of the zero-offset central ray on the reflector. All rays related to this normal wave (N-wave) are themselves

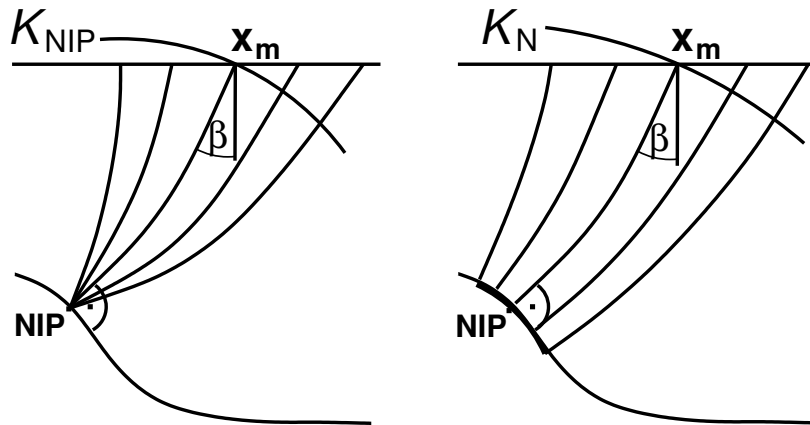


Figure 3.1: Hypothetical wavefronts (after Duvneek, 2004a): NIP-wave (left) and normal-wave (right). The wavefronts emerge at the surface at x_m with an angle β against the vertical and curvatures K_{NIP} and K_N , respectively.

normal to the interface at their incidence points. The matrix \mathbf{M} may be interpreted as the second horizontal traveltimes derivative matrix related to a wavefront which stems from a point source placed at the NIP on the reflector. Therefore, the corresponding wave is called NIP-wave. For a 2D sketch of both waves see Figure 3.1.

The symmetric 2×2 matrices \mathbf{K}_{NIP} and \mathbf{K}_N describe the curvature of the corresponding wavefronts at the emergence location x_m . Both wavefronts emerge with an angle β against the vertical in the azimuthal direction α .

Identifying CMP traveltimes described by \mathbf{M} with traveltimes along rays which all pass through the NIP of the ZO ray requires the so-called NIP-wave theorem (Chernyak and Gritsenko, 1979; Hubral, 1983) to be valid. The NIP-wave theorem states that up to second-order in the offset dimension traveltimes along common-reflection-point rays are identical to traveltimes along rays from a CMP family. Important to note at this point is that the NIP-wave theorem is an approximation. Its validity range strongly depends on the subsurface complexity. The traveltimes difference between rays from a CMP family and true CRP rays usually increases with increasing offset.

3.1.2 Practical aspects

The eight parameters in the 3D case or three parameters in the 2D case which define the CRS stacking operator are determined for each considered zero-offset sample by automatic coherence analysis. The used coherence criterion is semblance defined together with different other coherence measures in Taner and Koehler (1969) and Neidell and Taner (1971). Beside the stacked section/volume, this results in a number of additional output sections/volumes containing these best-fit attribute values and a section/volume containing the semblance values evaluated along the best-fit stacking operators. The attributes itself are meaningless in regions where no coherent reflection/diffraction energy is found in the prestack data. The physical interpretation in terms of kinematic wavefield attributes is only valid on coherent events characterised by relatively high semblance values. Thus, the coherence section/volume provides information where reflection events could be detected and reliable attribute values have been determined.

As these locations and the attribute values themselves are initially unknown a whole bunch of attribute values is tested for each output sample until the operator is found which best fits the reflection event in the prestack data. This states a multi-parameter optimisation problem for each ZO output sample. Sophisticated search strategies have been developed during the last years to make this problem computationally feasible. Throughout this thesis, the implementations described in [Mann \(2002\)](#) for the 2D case and in [Müller \(2003\)](#) for the 3D case are used. Both implementations split the search into several parts performed in subsets of the data. Firstly, matrix \mathbf{M} or its 2D equivalent is determined in CMP gathers/volumes and the data are stacked to yield a preliminary ZO section/volume. Therein, the remaining parameters are determined. The complete set of stacking parameters may then be subject to a simultaneous three (2D case) or eight (3D case) parameter optimisation. These optimised attributes are finally used to obtain the CRS stack section/volume.

The validity of the second-order traveltimes approximations (3.1) and (3.3) depends on the aperture chosen for the determination of the stacking parameters: it generally decreases with increasing distance in midpoint and offset direction from the considered ZO location. To obtain optimum results the apertures for the parameter search and stack have to be chosen carefully.

The best-fit stacking parameters themselves also depend on the chosen apertures. This effect has already been mentioned above and is known as spread-length bias. [Müller \(2006\)](#) presented a technique to correct the CRS stacking parameters for the spread-length bias influence. Especially, when the determined stacking parameters are used in further applications the aperture choice is of special importance. If the attributes are to be used in the tomographic inversion scheme presented below the aperture in the offset dimension has to be chosen in a way that the NIP-wave theorem is valid.

Beside the systematic effect of the spread-length bias, the stacking parameters are influenced by several other effects, e. g., varying signal-to-noise ratio and missing traces. All these effects result in unphysical fluctuations of the stacking parameters along continuous reflection events. These fluctuations may be removed by event-consistent smoothing strategies which are described in the following. Locally valid statistics allow to assess the uncertainty inherent in the kinematic wavefield attributes.

3.1.3 Event-consistent smoothing and uncertainty estimation for CRS stacking parameters

Theoretically, CRS attributes have two basic properties: they are virtually constant along the wavelet ([Mann and Höcht, 2003](#)) and they vary smoothly along the reflection event as this is related to an emerging continuous wavefront. Based on these observations, [Duvencck \(2004a\)](#) introduced an event-consistent smoothing strategy for the 2D case. This smoothing strategy helps to significantly improve CRS-based imaging results ([Hertweck et al., 2005](#)). I generalise this strategy to the 3D case in this section. The method may be applied to the kinematic wavefield attributes as well as to the corresponding first and second traveltimes derivatives. Both types of attributes ideally have above stated basic properties.

The implementation of the method is based on a small volume aligned with the reflection event in the ZO stacked data volume. Inside this event-aligned volume, locally valid statistics are applied to remove unphysical fluctuations and outliers from the stacking parameters. In time direction, the smoothing volume should not be larger than the wavelet of the considered event in order not to mix valuable information with noise or information related to different coherent events. In the spatial directions, it should be restricted to a few traces (depending on the midpoint spacing) in order to

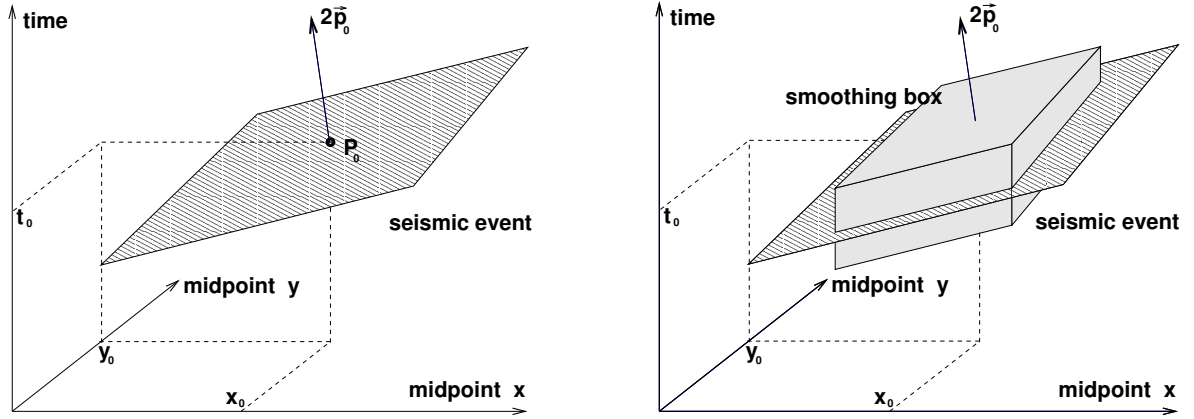


Figure 3.2: Event-aligned volume. Left: for each ZO sample P_0 the time dip is available from the CRS attributes as twice the horizontal slowness. Right: this time dip can be used to align a small volume with the reflection event around each considered sample.

avoid too strong variations of the wavefield attributes. In fact, the lateral size must be small enough to ensure locally almost constant attributes along the event. Otherwise, a bias might be introduced into the attribute values. In order to stay inside the considered reflection event, the smoothing volume is tilted according to the time dip of the reflection event in the stacked volume. This dip is given by twice the horizontal slowness \mathbf{p} as is obvious from equation (3.1). A schematic sketch of this event-aligned volume is depicted in Figure 3.2. Although a parallelepiped is shown a disk-shaped volume is used in practice.

Using equation (3.2a), one can easily determine the unit normal vector to the N- and NIP-wavefronts at \mathbf{x}_m :

$$\mathbf{n} = (\cos \alpha \sin \beta, \sin \alpha \sin \beta, \cos \beta)^T. \quad (3.9)$$

This can be done for each sample inside the aligned volume. In this way, the dip difference θ between wavefront normal vectors can be calculated from their inner product:

$$\theta = \arccos(\mathbf{n}_1 \cdot \mathbf{n}_2). \quad (3.10)$$

The dip difference θ gives an efficient criterion to distinguish between different events in the case of conflicting dip situations and, thus, may be used to prevent any mixing of information from different events.

All samples inside the event-aligned volume are subject to different rejection mechanisms and filtering processes: firstly, all samples with a dip difference θ beyond and/or a coherence below a user-defined threshold are removed. These samples are considered to be related to a different event or to be unreliable. The remaining samples are subject to a median filter to remove outliers: all attributes are ordered by magnitude and only a user-defined fraction around the median is selected. These selected attribute values are averaged to remove fluctuations. These steps are repeated automatically for all samples in the ZO volume. If there remains no sample inside the event-aligned volume after the filtering steps the original attribute values are kept. In this way, a set of smoothed attribute volumes is obtained. The procedure in the 2D case is analogous. However, the event-aligned volume reduces to a window.

Figure 3.3 shows a comparison between stacking velocities calculated from original and smoothed kinematic wavefield attributes according to equation (3.6) for a synthetic 3D example. The stacking

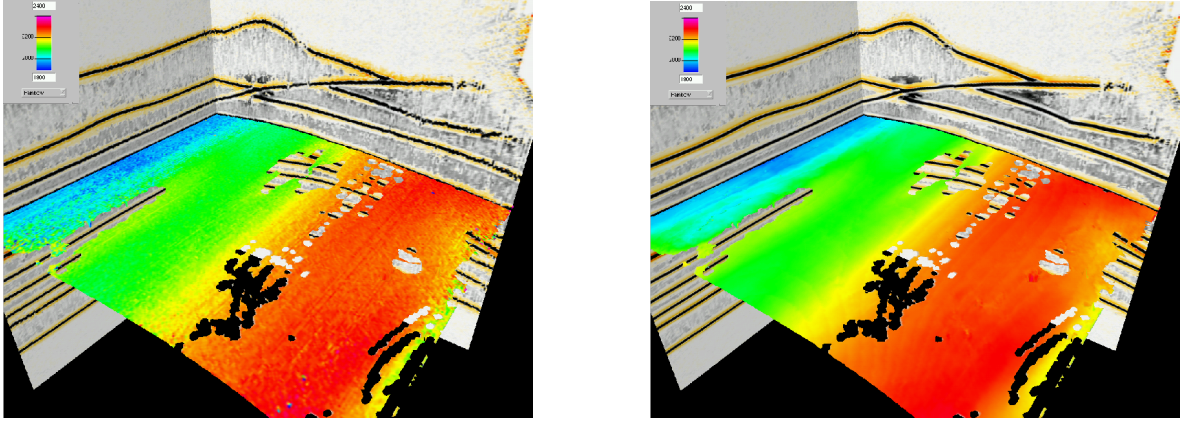


Figure 3.3: Stacking velocity calculated from original (left) and smoothed (right) kinematic wavefield attributes along an identified reflection event in a CRS stacked ZO volume. The stack itself has been obtained using original and smoothed attributes, respectively.

velocity values are shown for a selected reflection event identified in the CRS stack volume and cover a range between 1900 m/s and 2400 m/s. Virtually all fluctuations have been removed. Comparing the stack results obtained with smoothed and original attributes shown in two orthogonal directions, it is obvious that not only the attribute values themselves but also the stacked image is improved by above described event-consistent smoothing algorithm.

The stacked image obtained with smoothed attributes, although showing increased continuity of reflection events, in many cases comes along with a decreased coherence along the CRS stacking operator. This is due to the fact that the operator parameterised with smoothed attributes is no longer a best fit operator. However, an increased coherence along the CRS operator with smoothed attributes might occur at locations where the applied attribute search strategies in the CRS processing failed. At these locations, unreliable attribute values may be substituted with information from the local vicinity, provided that the dip information of the central sample is reasonable.

The above described strategy is suitable to pre-condition the input for the tomographic inversion scheme presented below. However, there is no possibility to assess the uncertainty of the stacking parameters by the above described method. For this purpose, I present a different technique. I restrict the description to the 2D case. Again, all samples inside the event-aligned window are used to estimate the smooth attribute values. Instead of using a combination of mean/median filtering, a linear model is fit to the attribute values minimising their absolute deviation from this model. The basic assumption is that the attribute values within the event-aligned window locally follow a linear trend.

Let y denote any of the CRS stacking parameters or kinematic wavefield attributes, respectively. Within the event-aligned window it is assumed that this attribute locally follows

$$y(x - x_c, t|a, b) = a(x_c, t) + b(x_c, t)(x - x_c), \quad (3.11)$$

where x_c denotes the lateral position of the central sample. The two coefficients a and b are functions of lateral position and time. These two parameters are estimated for each sample in the ZO section by minimising

$$\sum_{i=1}^n |y_i - a - b(x_i - x_c)|, \quad (3.12)$$

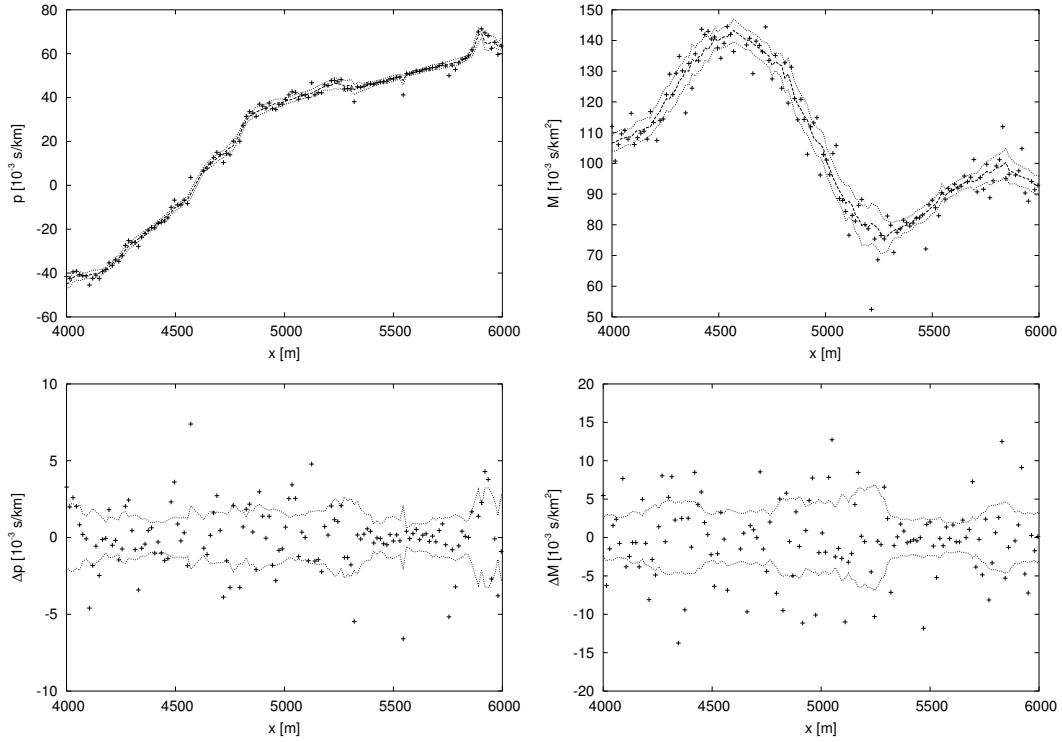


Figure 3.4: Attribute smoothing and uncertainty estimation using robust linear regression for the horizontal slowness (left-hand column) and the second traveltime derivative (right-hand column). The top row shows the attribute values before smoothing (crosses), after smoothing (dashed line), and the corresponding errors (dotted lines). The bottom row shows the same results centred around the smoothed attribute values to increase the display resolution.

where n denotes the number of samples within the event-aligned window. Like in the above described smoothing technique, all samples with a dip difference θ beyond and/or a coherence below a user-defined threshold are excluded. If only two or less samples remain after this rejection process the original attribute values are kept and the associated error is set to a pre-defined value which indicates that no smoothing has been performed. Otherwise, a robust linear regression is performed for each attribute type. The result are estimates of a and b for each sample in the ZO section. As the line is centred around x_c , the estimate of a directly provides the smoothed attribute value. Additionally, the mean absolute deviation from the linear model (3.11) is assigned as an uncertainty estimate for the considered attribute to the central sample. These uncertainty estimates are applied in the tomographic inversion scheme described below as individual weights for each input data point component.

The results of the above described technique for the slowness p and the second traveltime derivative M are shown in Figure 3.4 for a segment of the lowermost reflection event in the CRS stack section depicted in Figure 3.7 at the end of this chapter. In the lower row, the difference between the original attribute values and the corresponding smoothed values is shown. The mean absolute deviation defines a symmetric error corridor depicted as grey lines. The crosses denote the original attribute values at the corresponding lateral position. Except of outliers, which are less weighted in the robust parameter estimation, all attribute values lie within the error range. The estimated attribute values show largely reduced fluctuations. Thus, they may be considered as a good estimate of the true values which

should vary smoothly along the reflection event. The corresponding error gives a good measure of their uncertainty.

3.1.4 Second-order diffraction traveltimes in terms of CRS stacking parameters

The kinematic wavefield attributes determined by the CRS method may be used to parameterise a second-order approximation of the kinematic response of a diffraction point in depth. As described above, the stacking parameter \mathbf{M} is related to a hypothetical wave stemming from a point source in depth. This interpretation is based on the NIP-wave theorem. For a true diffractor in the subsurface one naturally obtains $\mathbf{M} = \mathbf{N}$. To obtain the approximate second-order diffraction traveltime operator the matrix \mathbf{N} in equation 3.1 is replaced by matrix \mathbf{M} :

$$t^2(\mathbf{x}_m + \Delta\mathbf{x}_m, \mathbf{h}) = (t_0 + 2\mathbf{p} \Delta\mathbf{x}_m)^2 + 2t_0 (\Delta\mathbf{x}_m^T \mathbf{M} \Delta\mathbf{x}_m + \mathbf{h}^T \mathbf{M} \mathbf{h}) . \quad (3.13)$$

This is justified as the matrix \mathbf{M} is independent of the reflector curvature as long as the NIP-wave theorem is valid.

The approximate diffraction traveltime operator is easily expressed in terms of kinematic wavefield attributes using relation 3.2. Reduced to the 2D case, the operator reads in this notation

$$t^2(x_m, h) = \left(t_0 + \frac{2 \sin \beta}{v_0} \Delta x_m \right)^2 + \frac{2t_0 \cos^2 \beta}{v_0} (K_{\text{NIP}} \Delta x_m^2 + K_{\text{NIP}} h^2) . \quad (3.14)$$

This operator can be used for an automatic, attribute-based time migration (Mann, 2002).

3.2 NIP-wave tomography

The previous section covered the extraction of parameters describing second-order approximations of diffraction traveltime operators from the seismic prestack data. In this section, I discuss the velocity model building technique based on these second-order diffraction traveltime functions. As the method explicitly uses kinematic information describing NIP-waves it is called NIP-wave tomography. Introduced by Duvencek and Hubral (2002), theoretical and various practical aspects are extensively described in Duvencek (2004a). I will review this work with special emphasis on the 3D case. I extend the method to azimuth-dependent second-order traveltime information. The data-driven uncertainty estimates for CRS attributes are incorporated as data covariance operators. A strategy for the automated extraction and pre-conditioning of input for the tomography from the CRS attribute sections/volumes is introduced.

This section starts with a short description of the relationship between NIP-waves and velocity models. The inverse problem to be solved in NIP-wave tomography is formulated. Model and data components are described followed by the introduction of data-driven covariance operators. Then, the automated picking strategy is presented. Finally, I describe the tomographic inversion algorithm and several of its theoretical and practical aspects.

3.2.1 NIP-waves and velocities

In migration-based velocity analysis a model is considered to be consistent with the seismic data if all reflection signals stemming from a common reflection point in the true subsurface are migrated to

the same position. This criterion is in accordance with the one used in focusing analysis: downward continuation of the seismic data in a consistent velocity model leads to focusing at zero offset and zero traveltime.

The specular rays for a CRP connect source and receiver locations at the surface with the CRP. They are geometrically identical to rays pertaining to a wavefront originated from a point source placed at the CRP. In the previous section this wave has been introduced as NIP-wave. The CRP coincides with the NIP of the ZO ray on the reflector. Based on this interpretation imaging consists of collecting all energy stemming from a CRP and assigning it to a common image point. Thus, in a consistent velocity model all NIP-wavefronts related to coherent reflection events focus at zero traveltime if they are propagated back into the subsurface. Assuming the NIP-wave theorem to be valid, kinematic information describing the NIP-wave may be extracted from the seismic prestack data. This information can be used for a second-order approximation of diffraction traveltimes. Thus, the focusing of NIP-wavefronts is equivalent to the focusing of energy spread along a second-order diffraction traveltime operator.

The concept of focusing NIP-waves has found several applications. [Hubral and Krey \(1980\)](#) presented algorithms for the construction of 2D and 3D models consisting of iso-velocity layers separated by smoothly curved interfaces from stacking velocities in a layer-by-layer inversion. [Söllner and Yang \(2002\)](#) used visual focusing analysis of energy spread along second-order diffraction traveltime operators in a poststack depth migration velocity analysis loop.

The diffraction traveltime formula [3.13](#) using kinematic information related to a NIP-wave describes a second-order approximation of the true diffraction traveltimes. In general, this limits the applicability of NIP-wave tomography to subsurface models in which the kinematic wavefield attributes may be determined in a way that the NIP-wave theorem holds. In any case, the velocity models obtained with the tomographic inversion scheme have to be considered as consistent with the seismic data up to second order only. In more complex situations they are excellent initial models for a further update. Such an updating technique is presented in [Chapter 4](#).

Beside its approximate character, NIP-wave tomography has several advantages: no interfaces have to be introduced in the model description as each NIP-wave is related to a point source and, thus, can be treated independently of all others. This allows the usage of a smooth velocity model parameterisation and further implies that no continuous reflection events need to be identified in the prestack data like in conventional reflection traveltime tomography. The input for the inversion scheme is extracted in poststack ZO sections/volumes with a considerably increased signal-to-noise ratio. This extraction may be performed in a highly automated manner using the coherence information obtained during the stacking procedure. There is not even a need for the existence of continuous reflection events in the poststack section/volume as long as locally coherent reflection events exist and their associated kinematic wavefield attributes are reliable.

3.2.2 The inverse problem

All methods aiming at resolving the subsurface velocity distribution from geophysical measurements at the surface state an inverse problem. NIP-wave tomography is no exception to this rule. The inverse problem to be solved can be formulated in the following way: a velocity model needs to be found wherein all considered NIP-waves focus at zero traveltime if they are back-propagated into the subsurface.

Back-propagation of NIP-waves is easily performed for any velocity model as the starting position and initial propagation direction are readily available from the CRS stack attributes. The task is to find a velocity model in which all considered NIP-waves focus at zero traveltimes, i. e., the elements of the wavefront curvature matrix \mathbf{K}_{NIP} vanish for $t = 0$. However, solving the inverse problem in this manner implies to fix the kinematic properties describing the NIP-waves and, thus, to neglect all errors inherent in the measured data. The kinematic wavefield attributes or CRS stacking parameters, respectively, are affected by noise or measurement errors. A method to quantify this error has been presented in Section 3.1.3.

In the implementation the NIP-waves are not back-propagated but the quantities describing their kinematic properties are forward modelled and compared to the corresponding quantities extracted from the seismic prestack data by means of the CRS stack method. The tomographic inversion scheme aims at minimising the difference between the forward-modelled and measured data. Thus, it accounts for potential measurement errors. A similar approach has been used by [Billette and Lambaré \(1998\)](#) in the context of stereotomography. The inverse problem is formulated as the minimisation of a cost-function in the least-squares sense. For a detailed description of the procedure see Chapter 5 and [Duvencq \(2004a\)](#).

The inverse problem is non-linear as the forward-modelled data are non-linear functions of the velocity model. Consequently, it is solved iteratively computing a model update in each iteration. The solution as well as the regularisation and additional constraints that could be applied are described in Chapter 5 in a way valid for both inversion methods presented in this thesis.

3.2.3 Data and model components

As explained above, each NIP-wave can be treated independently. Let me assume that n_{data} NIP-waves have been selected to be used in the tomographic inversion. For each of these NIP-waves the quantities describing their kinematic properties are used as input data for the tomography. These quantities are extracted from the corresponding sections/volumes of kinematic wavefield attributes which are output of the CRS stack method. An efficient automated strategy to select reliable pick locations in the CRS stack results is presented below. For each selected pick associated with an individual NIP-wave the normal ray emergence location \mathbf{x}_m , the one-way traveltimes T as well as the first and second traveltimes derivatives are extracted automatically from the corresponding attribute volumes. Each NIP-wave is represented by

$$(T, \mathbf{x}_m, \mathbf{p}, \mathbf{M}) \quad (3.15)$$

in the 3D case with azimuth-dependent second traveltimes derivatives. The quantities in (3.15) make up a single data point. In the case of seismic prestack data, which only allow the determination of the second traveltimes derivative in a certain azimuth, the scalar M_ϕ defined in equation (3.7) replaces the matrix \mathbf{M} in the data point description. All data points are summarised in the vector \mathbf{d}_{mea} of measured data which is symbolically expressed by

$$\mathbf{d}_{\text{mea}} = (T_i, \mathbf{x}_{m_i}, \mathbf{p}_i, \mathbf{M}_i) , \quad i = 1, \dots, n_{\text{data}} . \quad (3.16)$$

In the 2D case, the position vector, the horizontal slowness vector, and the matrix of second traveltimes derivatives reduce to scalar quantities.

The implementation of the inversion scheme is based on a global Cartesian coordinate system. The x-axis and y-axis coincide with the axes defining the horizontal two-dimensional Cartesian coordinate

system attached to the measurement surface which is used in the description of the prestack data geometry (see Chapter 1). The positive z -axis is pointing upwards.

Associated with each NIP-wave is its focusing point in depth, the NIP of the ZO ray on the reflector. As the NIP-waves are not back-propagated (which would allow to calculate the NIP in the current model) but forward modelled, the normal incidence point as well as the initial normal ray direction are part of the model description. They are considered to be unknown and are inverted for. The initial direction of the normal ray does not need to be included completely: in the 2D case only one quantity and in the 3D case two quantities describing that direction have to be used. The remaining component of the vector describing the initial ray direction is not independent as the ray has to satisfy the eikonal equation (2.12). For each single data point in the 3D case the quantities

$$(x, y, z, e_x, e_y) \quad (3.17)$$

are part of the model description where e_x and e_y denote the two horizontal components of the unit tangent vector to the normal ray at the NIP. In the 2D case, the take-off angle Θ of the normal ray at the NIP is used. The components of the model description related to the NIP in the 2D case are given by

$$(x, z, \Theta) . \quad (3.18)$$

For all NIP locations applies $z < 0$ m as the measurement surface is assumed to coincide with $z = 0$ m. The take-off angle Θ of the ZO ray in the 2D case coincides with the local reflector dip as the ZO ray is normal to the reflector at the NIP. This angle is positive if the horizontal slowness component of the normal ray at the NIP is positive and vice versa, i. e., $0 \leq |\Theta| \leq \pi$. It is measured from the normal ray to the positive z -axis. The same convention applies to the angle enclosed by the reflector normal and the z -axis in the 3D case. However, the possible reflector dips are restricted to the range $0 \leq |\Theta| < \pi/2$ because the implementation of the forward modelling is based on the reduced Hamiltonian in Cartesian coordinates (see Chapter 2). This allows no turning rays with respect to the z coordinate. The azimuthal direction of the projection onto the horizontal plane of the reflector normal is measured counterclockwise starting at the positive x -axis.

Together with the B-spline coefficients of the velocity model parameterisation (see Appendix A) the above quantities related to the NIPs of the ZO rays build the model vector \mathbf{m} . Symbolically expressed for the 3D case, it reads

$$\mathbf{m} = (x_i, y_i, z_i, e_{x_i}, e_{y_i}, v_{jkl}) , \quad (3.19)$$

with $i = 1, \dots, n_{\text{data}}$, $j = 1, \dots, n_x$, $k = 1, \dots, n_y$, and $l = 1, \dots, n_z$. The number of node points defining the grid for the B-spline coefficients are given by n_x , n_y , and n_z for the corresponding spatial directions. The model vector for the 2D case reads

$$\mathbf{m} = (x_i, z_i, \Theta_i, v_{jk}) , \quad (3.20)$$

with $i = 1, \dots, n_{\text{data}}$, $j = 1, \dots, n_x$, and $k = 1, \dots, n_z$.

3.2.4 Data-driven determination of covariance operators

The above derived data and model vectors are those used in the minimisation of the cost-function Ψ defined in Chapter 5. Without regularisation term this cost-function reads

$$\Psi = \|\mathbf{d}_{\text{mea}} - \mathbf{d}_{\text{mod}}\|^2 = \frac{1}{2} \Delta \mathbf{d}^T(\mathbf{m}) \mathbf{C}_D^{-1} \Delta \mathbf{d}(\mathbf{m}) . \quad (3.21)$$

Matrix C_D effectively weights the different data points. In the original implementation by [Duvencek \(2004a\)](#) the matrix is used as a diagonal matrix weighting the different data classes, i. e., each data class like traveltime T has its own constant weight. This is necessary to balance the different data types which all have different physical and numerical dimensions. I extend the inversion scheme such that each individual data value may have its own weight. These weights are directly determined from the CRS attribute sections/volumes using the method presented in Section 3.1.3.

The uncertainty estimation procedure introduced in Section 3.1.3 does not yield any error estimates for the location x_m and traveltime T . The errors for this part of the input data have to be determined in a different way. Meaningful error estimates may be found directly from the data geometry: the picking strategy extracts the input for the tomography from the discrete output sections/volumes of the CRS stack. Therefore, I consider half of the time sampling and half of the mean midpoint spacing as reasonable error estimates for traveltime T and location x_m .

The above introduced weights for traveltime T and x_m together with the data-driven uncertainties for the CRS stacking parameters are the elements of the matrix C_D which can then be interpreted as a data covariance matrix. It is a diagonal matrix as no correlations between the individual data vector components are assumed. In order to account for the variable numerical dimensions of the different data types each data vector component is additionally weighted with a constant weight according to its type. Thus, the influence of the different data classes on the inversion result is controlled by their corresponding constant weight whereas the influence of each single pick is controlled by its data-driven uncertainty. In the data example presented at the end of this chapter, the average of all individual errors belonging to the same data class is used as constant weight for that data class. This is sufficient for numerical balancing and does not require any user interaction. The constant weights for each data type are denoted by w_T , w_M , w_p , and w_x , respectively. This notation applies for the 2D as well as for the 3D case.

3.2.5 Automated picking strategy

As already outlined above, each set of quantities describing the kinematic properties of a NIP-wave is considered as a single data point in the tomographic inversion. As the NIP-wave is related to a point source, there is no need to extract the input for the tomography along identified continuous reflection events in the stacked ZO section/volume. The identification of locally coherent events is sufficient. This identification can be highly automated based on the CRS stacking parameters and their optionally calculated uncertainties. For the decision whether a sample in the ZO section/volume is located on a locally coherent event and, thus, is a valid pick location, the idea of the event-aligned window is used again. If the considered sample is part of a locally coherent event criteria similar to those used in the event-consistent smoothing apply to the samples inside the event-aligned volume. This allows to formulate an algorithm for the automatic identification of valid pick locations and the extraction of the input for the tomography from the CRS results.

For each trace in the ZO section/volume

- search the coherence maximum on the selected trace and go to the nearest maximum of the stack envelope,
- optionally, check whether the selected sample contains attribute values with standard errors smaller than a user-defined threshold,

- optionally, check whether the amplitude of the central sample exceeds a user-defined threshold,
- align a volume with the reflection event using the first traveltimes derivatives,
- check whether a user-defined percentage of all samples inside the window
 - has coherence values higher than a given threshold and
 - has a dip difference θ below a given threshold with respect to the central sample,
- optionally, check whether a user-defined percentage of all samples inside the window
 - contains attribute values with associated standard errors below a user-defined threshold,
- continue on the selected trace until a user-defined maximum number of picks on this trace is found.

Valid picks are not only selected according to their coherence value. Additionally, information from neighbouring samples on the same reflection event is taken into account. This allows to check whether the pick location under consideration is actually part of a locally coherent reflection event. Considering uncertainties of the kinematic wavefield attributes already in the picking stage allows to extract the most reliable information from the CRS stack results. The pick locations are shifted to the maximum of the stack envelope to align all picks on the same reflection event at a consistent time level.

The above described picking strategy works without any human intervention and is very fast. However, not all selected pick locations contain meaningful input for the tomography. In principle, all locally coherent events related to different wave types and converted wave modes as well as coherent noise may be picked. Therefore, an additional editing step is necessary in which all reliable picks related to P-P primary reflection events are extracted from the initial set of picks. A convenient way for this purpose is to use constraints on the stacking velocity range. The stacking velocity can be calculated for each pick from the extracted attribute values using relationship (3.6) in the 3D case or equation (3.8) in the 2D case.

However, editing the automatically selected picks might not be sufficient to remove all outliers. Especially in complex data, the only way to identify reliable reflection events is manual picking. In that case, only interpreted events contribute in the tomographic inversion. Manual picking of input for NIP-wave tomography is relatively easy compared to the picking necessary for other tomographic inversion schemes as it is performed in the poststack domain. The considerably increased signal-to-noise ratio after application of the CRS stack method makes the interpretation of reflection events much easier than in the prestack data.

3.2.6 The tomographic inversion algorithm

The minimisation of the cost-function Ψ defined in Chapter 5 is performed in an iterative way. Assuming the appropriate input for the tomography has been extracted from the seismic prestack data, the algorithm proceeds as follows:

- An initial velocity model is set up by defining the B-spline node sequence in each spatial direction and assigning appropriate values to the B-spline coefficients. The components of the data covariance matrix C_D are defined by either using constant weights for each data class or the combination of these constant weights with the data-derived uncertainties. Additional constraints and appropriate weights for the regularisation need to be specified.

- Initial NIP locations and initial normal ray directions are computed in the initial velocity model. A normal ray is traced down into the subsurface until the one-way time is depleted. Its starting position and direction are part of the input data. The ray end point defines the initial NIP location. The ray slowness vector at the NIP is used to calculate the reflector normal.
- The initial data vector \mathbf{d}_{mea} is forward modelled using dynamic ray tracing. Simultaneously, the Fréchet derivatives, i. e., the elements of matrix \mathbf{F} (see Chapter 5), are calculated.
- The cost-function Ψ is evaluated calculating the data misfit, the regularisation, as well as additional constraints and their associated misfit.
- The linear system (5.16) is solved using the LSQR (Paige and Saunders, 1982a,b) algorithm to obtain the model update vector $\Delta\mathbf{m}$. Details follow below and are described in Chapter 5. Each iteration starts with this step.
- The model update vector is multiplied by a factor $0 < \lambda \leq 1$ to obtain the new model. Optionally, the regularisation weights are decreased to allow for more model details in the next iteration. The cost-function is recalculated computing the data misfit and additional constraints in the new model.
- If the cost-function has increased, the updated model is rejected, λ is decreased, and the previous step is repeated until the cost-function decreases or λ falls below a certain threshold. In the latter case, it is assumed that a minimum of the cost-function has been reached and the inversion is terminated.
- If the cost-function has decreased in the updated model the Fréchet derivatives are calculated and the next iteration is started with the inversion of system (5.16).

Except for the initial set up, all steps are performed without human interaction.

The regularisation weights may be decreased from iteration to iteration based on the decrease of the cost-function. This allows to resolve more and more details in the velocity distribution with decreasing data misfit. However, such a strategy is difficult to control. Therefore, a user-driven strategy might be applied: based on an examination of the results after each iteration the inversion may be stopped and the regularisation weights decreased. The inversion is then restarted from the current model.

Forward modelling and calculation of Fréchet derivatives

Forward modelling and the calculation of the Fréchet derivatives is based on ray and ray perturbation theory as it is described in Chapter 2. In each iteration of the inversion algorithm outlined above, the quantities defining a single data point have to be forward modelled for each pick. In the 3D case with azimuth-dependent second traveltimes derivatives, these are

$$(\mathbf{T}, \mathbf{x}_m, \mathbf{p}, \mathbf{M})_i, \quad i = 1, \dots, n_{\text{data}} .$$

In the case of narrow azimuth data, where the second traveltimes derivative can be extracted from the seismic prestack data only for a selected azimuth direction, the matrix \mathbf{M} is replaced by M_ϕ defined in equation (3.7). The corresponding quantities in the 2D case are

$$(T, x_m, p, M)_i, \quad i = 1, \dots, n_{\text{data}} .$$

To avoid too many indices in this chapter I have changed the notation: (x_1, x_2, x_3) in Chapter 2 corresponds to (x, y, z) .

The emergence location and horizontal slowness in the 2D case are computed solving the kinematic ray tracing system (2.14) in the global Cartesian coordinate system with arclength s being the independent variable. In the 2D case the y -component vanishes. Each normal ray to be traced starts at the corresponding NIP location in the current model and ends at the surface. The initial ray slowness vector at the NIP is given by $(\sin \Theta / v_{\text{NIP}}, \cos \Theta / v_{\text{NIP}})^T$ where v_{NIP} denotes the velocity at the NIP. Traveltimes T is obtained integrating equation (2.15) along the normal ray. The second traveltimes derivative M is forward modelled using dynamic ray tracing. In the 2D case, dynamic ray tracing is performed in ray-centred coordinates. Solving (2.71) for initial point source and initial plane wave conditions along the normal ray yields the second traveltimes derivative in ray-centred coordinates due to a point source placed at the NIP. According to equation 2.74 it reads

$$M^{(q)} = \frac{P_2}{Q_2} .$$

Assuming the near-surface velocity to be constant within the paraxial region around the normal ray emergence location, the transformation of this expression to Cartesian coordinates at the measurement surface yields

$$M = M^{(q)} \cos^2 \beta ,$$

where β denotes the emergence angle of the normal ray. This angle is easily determined from the slowness vector and the velocity at the surface.

The 3D implementation is based on kinematic and dynamic ray tracing in Cartesian coordinates using the reduced systems (2.17) and (2.32) with z being the integration variable along the normal ray. The application of the reduced systems in Cartesian coordinates avoids many B-spline evaluations: all derivatives of the velocity with respect to the z -coordinate which would be necessary, e. g., in ray-centred coordinates do not appear. For this reason the application of the reduced systems is supposed to speed up the inversion. However, as z is the independent variable no turning rays with respect to the z -coordinate may be handled. The normal ray emergence location and the horizontal slowness components are directly obtained integrating (2.17) from the NIP to the surface. Traveltimes T is determined integrating (2.19) along the normal ray. The matrix of second traveltimes derivatives is directly obtained from the solution of the dynamic ray tracing system (2.32). No additional transformations are required. The matrix \mathbf{M} is given by

$$\mathbf{M} = \mathbf{P}_2 \mathbf{Q}_2^{-1} ,$$

where \mathbf{P}_2 and \mathbf{Q}_2 are solutions to system (2.32) for initial point source conditions. In the case of narrow azimuth seismic data acquisition where only the component M_ϕ can be extracted from the seismic prestack data (see Section 3.1) the corresponding forward-modelled second traveltimes derivative is calculated from the matrix \mathbf{M} according to

$$M_\phi = \check{\mathbf{e}}_\phi^T \mathbf{M} \check{\mathbf{e}}_\phi .$$

The azimuthal direction is given by the vector $\check{\mathbf{e}}_\phi$. The ray tracing systems used in the 2D and in the 3D case are integrated using second-order or fourth-order Runge-Kutta schemes (see, e. g., Press et al., 1992).

The Fréchet derivatives are the first partial derivatives of the data components with respect to the model components. For the general 3D case with azimuth-dependent second-order traveltime derivatives, they are symbolically written as

$$\frac{\partial (T_i, \mathbf{p}_i, \mathbf{M}_i, \mathbf{x}_{m_i})}{\partial (x_i, y_i, z_i, e_{x_i}, e_{y_i}, v_{jkl})}, \quad (3.22)$$

where $i = 1, \dots, n_{\text{data}}$, $j = 1, \dots, n_x$, $k = 1, \dots, n_y$, and $l = 1, \dots, n_z$. Similar expressions may be given for the 3D case with narrow azimuth information and the 2D case. The actual computation of the Fréchet derivatives is outlined in Appendix B. It is based on ray perturbation theory in ray-centred coordinates in the 2D case and reduced Cartesian coordinates in the 3D case.

The partial derivatives of a selected data point vanish for all model components not related to the corresponding NIP location as each pick is treated independently. Thus, the matrix \mathbf{F} containing the Fréchet derivatives is sparse. The structure of this matrix is described in the following for the most general case of 3D data with azimuth-dependent second traveltime derivatives and individual weights for each pick.

The structure of the tomographic matrix

The above outlined sparseness of the Fréchet derivative matrix is taken into account in the implementation by keeping only non-zero elements in the memory. The tomographic matrix $\hat{\mathbf{F}}$ involved in the computation of the model update vector $\Delta \mathbf{m}$ contains the matrix of Fréchet derivatives as upper left sub-matrix. For the general 3D case it has the following form:

$$\hat{\mathbf{F}} = \begin{pmatrix} \frac{w_T}{\sigma_{T_i}} \frac{\partial T_i}{\partial x_j} \delta_{ij} & \frac{w_T}{\sigma_{T_i}} \frac{\partial T_i}{\partial y_j} \delta_{ij} & \frac{w_T}{\sigma_{T_i}} \frac{\partial T_i}{\partial z_j} \delta_{ij} & \frac{w_T}{\sigma_{T_i}} \frac{\partial T_i}{\partial e_{x_j}} \delta_{ij} & \frac{w_T}{\sigma_{T_i}} \frac{\partial T_i}{\partial e_{y_j}} \delta_{ij} & \frac{w_T}{\sigma_{T_i}} \frac{\partial T_i}{\partial v_{jkl}} \\ \frac{w_M}{\sigma_{M_{xx_i}}} \frac{\partial M_{xx_i}}{\partial x_j} \delta_{ij} & \frac{w_M}{\sigma_{M_{xx_i}}} \frac{\partial M_{xx_i}}{\partial y_j} \delta_{ij} & \frac{w_M}{\sigma_{M_{xx_i}}} \frac{\partial M_{xx_i}}{\partial z_j} \delta_{ij} & \frac{w_M}{\sigma_{M_{xx_i}}} \frac{\partial M_{xx_i}}{\partial e_{x_j}} \delta_{ij} & \frac{w_M}{\sigma_{M_{xx_i}}} \frac{\partial M_{xx_i}}{\partial e_{y_j}} \delta_{ij} & \frac{w_M}{\sigma_{M_{xx_i}}} \frac{\partial M_{xx_i}}{\partial v_{jkl}} \\ \frac{w_M}{\sigma_{M_{xy_i}}} \frac{\partial M_{xy_i}}{\partial x_j} \delta_{ij} & \frac{w_M}{\sigma_{M_{xy_i}}} \frac{\partial M_{xy_i}}{\partial y_j} \delta_{ij} & \frac{w_M}{\sigma_{M_{xy_i}}} \frac{\partial M_{xy_i}}{\partial z_j} \delta_{ij} & \frac{w_M}{\sigma_{M_{xy_i}}} \frac{\partial M_{xy_i}}{\partial e_{x_j}} \delta_{ij} & \frac{w_M}{\sigma_{M_{xy_i}}} \frac{\partial M_{xy_i}}{\partial e_{y_j}} \delta_{ij} & \frac{w_M}{\sigma_{M_{xy_i}}} \frac{\partial M_{xy_i}}{\partial v_{jkl}} \\ \frac{w_M}{\sigma_{M_{yy_i}}} \frac{\partial M_{yy_i}}{\partial x_j} \delta_{ij} & \frac{w_M}{\sigma_{M_{yy_i}}} \frac{\partial M_{yy_i}}{\partial y_j} \delta_{ij} & \frac{w_M}{\sigma_{M_{yy_i}}} \frac{\partial M_{yy_i}}{\partial z_j} \delta_{ij} & \frac{w_M}{\sigma_{M_{yy_i}}} \frac{\partial M_{yy_i}}{\partial e_{x_j}} \delta_{ij} & \frac{w_M}{\sigma_{M_{yy_i}}} \frac{\partial M_{yy_i}}{\partial e_{y_j}} \delta_{ij} & \frac{w_M}{\sigma_{M_{yy_i}}} \frac{\partial M_{yy_i}}{\partial v_{jkl}} \\ \frac{w_p}{\sigma_{p_{x_i}}} \frac{\partial p_{x_i}}{\partial x_j} \delta_{ij} & \frac{w_p}{\sigma_{p_{x_i}}} \frac{\partial p_{x_i}}{\partial y_j} \delta_{ij} & \frac{w_p}{\sigma_{p_{x_i}}} \frac{\partial p_{x_i}}{\partial z_j} \delta_{ij} & \frac{w_p}{\sigma_{p_{x_i}}} \frac{\partial p_{x_i}}{\partial e_{x_j}} \delta_{ij} & \frac{w_p}{\sigma_{p_{x_i}}} \frac{\partial p_{x_i}}{\partial e_{y_j}} \delta_{ij} & \frac{w_p}{\sigma_{p_{x_i}}} \frac{\partial p_{x_i}}{\partial v_{jkl}} \\ \frac{w_p}{\sigma_{p_{y_i}}} \frac{\partial p_{y_i}}{\partial x_j} \delta_{ij} & \frac{w_p}{\sigma_{p_{y_i}}} \frac{\partial p_{y_i}}{\partial y_j} \delta_{ij} & \frac{w_p}{\sigma_{p_{y_i}}} \frac{\partial p_{y_i}}{\partial z_j} \delta_{ij} & \frac{w_p}{\sigma_{p_{y_i}}} \frac{\partial p_{y_i}}{\partial e_{x_j}} \delta_{ij} & \frac{w_p}{\sigma_{p_{y_i}}} \frac{\partial p_{y_i}}{\partial e_{y_j}} \delta_{ij} & \frac{w_p}{\sigma_{p_{y_i}}} \frac{\partial p_{y_i}}{\partial v_{jkl}} \\ \frac{w_x}{\sigma_{x_{m_i}}} \frac{\partial x_{m_i}}{\partial x_j} \delta_{ij} & \frac{w_x}{\sigma_{x_{m_i}}} \frac{\partial x_{m_i}}{\partial y_j} \delta_{ij} & \frac{w_x}{\sigma_{x_{m_i}}} \frac{\partial x_{m_i}}{\partial z_j} \delta_{ij} & \frac{w_x}{\sigma_{x_{m_i}}} \frac{\partial x_{m_i}}{\partial e_{x_j}} \delta_{ij} & \frac{w_x}{\sigma_{x_{m_i}}} \frac{\partial x_{m_i}}{\partial e_{y_j}} \delta_{ij} & \frac{w_x}{\sigma_{x_{m_i}}} \frac{\partial x_{m_i}}{\partial v_{jkl}} \\ \frac{w_x}{\sigma_{y_{m_i}}} \frac{\partial y_{m_i}}{\partial x_j} \delta_{ij} & \frac{w_x}{\sigma_{y_{m_i}}} \frac{\partial y_{m_i}}{\partial y_j} \delta_{ij} & \frac{w_x}{\sigma_{y_{m_i}}} \frac{\partial y_{m_i}}{\partial z_j} \delta_{ij} & \frac{w_x}{\sigma_{y_{m_i}}} \frac{\partial y_{m_i}}{\partial e_{x_j}} \delta_{ij} & \frac{w_x}{\sigma_{y_{m_i}}} \frac{\partial y_{m_i}}{\partial e_{y_j}} \delta_{ij} & \frac{w_x}{\sigma_{y_{m_i}}} \frac{\partial y_{m_i}}{\partial v_{jkl}} \\ \mathbf{0}_v & & & & & w_v \frac{\partial v^{a-priori}}{\partial v_{jkl}} \\ \mathbf{0}_{v_q} & & & & & w_{v_q} \frac{\partial |\nabla_{\perp} v|_j}{\partial v_{jkl}} \\ \mathbf{0}_B & & & & & \mathbf{B} \end{pmatrix}, \quad (3.23)$$

with

$$\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

The corresponding matrices for the 2D case as well as for the 3D case with limited azimuth second-order traveltime derivative M_ϕ have analogous structures.

The individual sub-matrices of (3.23) have different geometries. All matrix dimensions given in the following have the format ‘number of rows \times number of columns’. Each sub-matrix containing partial derivatives of a certain data class like traveltimes T with respect to the NIP related part of the model description is of size $n_{\text{data}} \times n_{\text{data}}$, the sub-matrices containing the corresponding partial derivatives with respect to the B-spline coefficients each have a size of $n_{\text{data}} \times n_x n_y n_z$. The partial derivatives of the additional constraints (described in Chapter 5) with respect to the NIP-related model components vanish. The corresponding sub-matrices $\mathbf{0}_v$ and $\mathbf{0}_{vq}$ contain $n_{\text{vdata}} \times 5n_{\text{data}}$ and $n_{\text{data}} \times 5n_{\text{data}}$ elements, respectively. The remaining sub-matrices for the additional constraints containing the derivatives with respect to the B-spline coefficients have sizes of $n_{\text{vdata}} \times n_x n_y n_z$ and $n_{\text{vqdata}} \times n_x n_y n_z$, respectively. The remaining non-zero part is built by matrix \mathbf{B} related to the regularisation (see Chapter 5) and has a size of $n_x n_y n_z \times n_x n_y n_z$. As the regularisation is independent of the NIP-related part of the model description another matrix $\mathbf{0}_B$ containing only zeros is needed. It contains $n_x n_y n_z \times 5n_{\text{data}}$ elements. The meaning and derivation of the various weighting factors is described in Section 3.2.4.

Parallelisation

The tomographic matrix is very large for common seismic survey sizes, especially in the 3D case. Depending on the number of picks and B-spline nodes the tomographic system soon becomes too large to be solved on a single CPU system. Although the run-time decreases significantly, the main reason for a parallelised implementation is the lack of sufficient memory. The applied parallelisation strategy is simple: as all picks are treated independently, they may be distributed on several CPUs. On each CPU the corresponding forward modelling and calculation of Fréchet derivatives is performed. The Fréchet derivatives and the data misfit related to local picks are only kept in the local memory. This is an efficient strategy to reduce memory requirements and run-time. However, the LSQR algorithm (Paige and Saunders, 1982a,b) used for the computation of the model update vector $\Delta \mathbf{m}$ had to be adopted: as the matrices and vectors involved in internal matrix-vector products are distributed among several CPUs, these matrix-vector products had to be parallelised as well. The parallelisation of the tomographic inversion code is based on the message passing interface (MPI).

3.3 A velocity model building work flow based on second-order diffraction traveltimes functions

The scheme for velocity model building based on analytic second-order approximations of diffraction traveltimes functions combines all aspects described in the previous sections. I summarise all this information in a work flow for velocity model estimation which starts with the application of the CRS stack method and ends with NIP-wave tomography. This work flow is depicted in Figure 3.5.

The first step is the application of the CRS stack method. Throughout this thesis, I use the implementations described in Mann (2002) for the 2D case and in Müller (2003) for the 3D case. Beside the simulated ZO stack result itself, the CRS stack yields a section/volume for each of the stacking parameters or kinematic wavefield attributes, respectively. Additional output is a section/volume containing the coherence along the CRS stack operator for each ZO sample.

Optional, but very useful, is the application of the event-consistent smoothing strategy to further improve the CRS imaging result. More important for velocity model building is its character as a

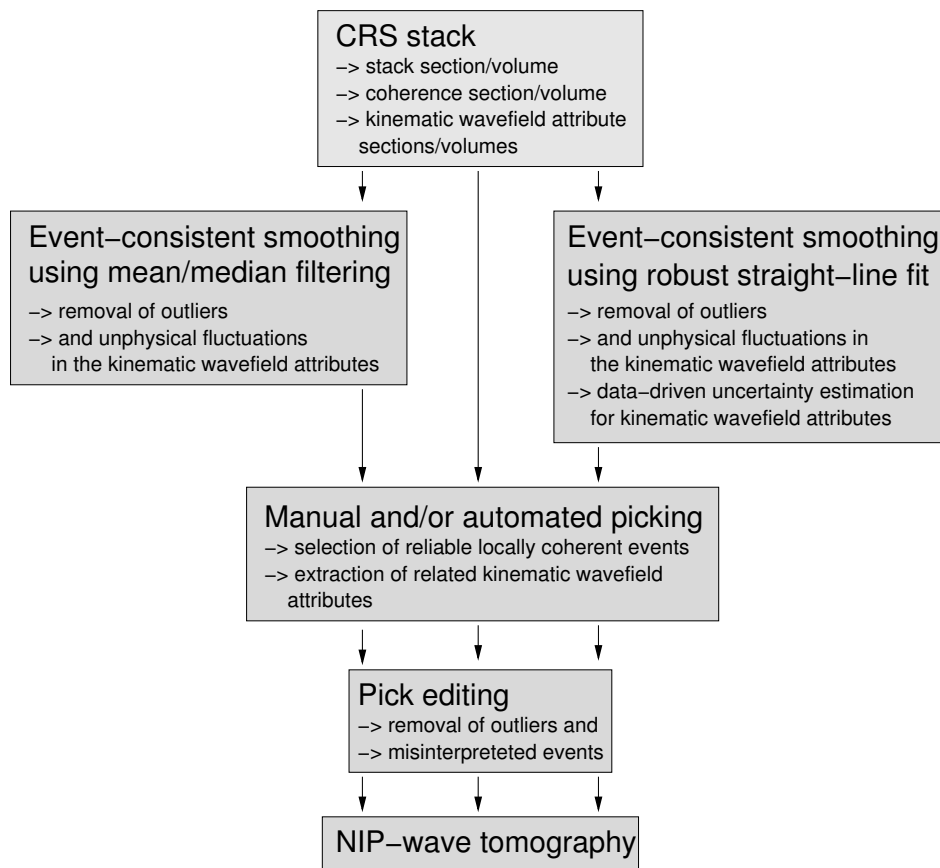


Figure 3.5: The velocity model building work flow based on second-order approximations of diffraction traveltimes functions. Three different paths incorporating different kinds of input data pre-conditioning are indicated with arrows.

pre-conditioning tool for the input of NIP-wave tomography. The application of the event-consistent smoothing based on a combination of mean and median filtering does not allow to properly reflect the uncertainties inherent in the attribute determination in the tomographic inversion. Only constant weights for different data classes can be applied. The event-consistent smoothing by means of a robust local linear regression allows to estimate the uncertainties of the CRS stacking parameters or kinematic wavefield attributes and, thus, the computation of data-driven covariance operators for the tomographic inversion.

The next step is the selection of reliable pick locations and the extraction of the input information necessary for the tomography from the CRS results. This can either be done by manual picking or by the automated strategy presented in Section 3.2.5. In the latter case, the picks should be carefully examined in order to remove outliers due to, for example, locally coherent events related to multiply reflected energy. One possibility for outlier identification is the visualisation of the picks as stacking velocity (calculated from the kinematic information) versus traveltimes.

With such prepared input, the tomographic inversion can be started with the definition of the initial model. The tomographic inversion algorithm itself proceeds automatically as described in Section 3.2.6.

3.4 Data example

The above presented work flow for velocity model building based on second-order diffraction traveltimes operators is demonstrated in this section using a 2D synthetic data example. For a 3D data example I refer to Chapter 6. The complete work flow described in the previous section is applied including the estimation of uncertainties for the CRS stacking parameters and their application as data covariance operators in the tomographic inversion algorithm.

3.4.1 Data description

The data consists of primary reflection events which have been forward modelled using a wavefront construction technique. The maximum offset in the prestack data is 4000 m. The complete data set consists of 53333 traces, each with 1001 samples. The sampling interval is 4 ms. The traces are distributed on 933 CMP gathers with a midpoint spacing of 15 m. A zero-phase Ricker wavelet with a peak frequency of 30 Hz has been used as source signal. For the application in this section, uniformly distributed random noise, band-limited to the frequency range of the source signal, has been added to the raw data. The signal-to-noise ratio has been chosen low enough to ensure fluctuations of the CRS stacking parameters. Actually, the noise level is so high that most events can hardly be detected with the naked eye. This is demonstrated in Figure 3.6 where a selected CMP gather is shown with and without noise.

In this example, only a maximum offset of 2000 m is used. This is the range wherein I assume the second-order approximation to be valid in the sense that the kinematic wavefield attributes actually describe kinematic properties of NIP-waves.

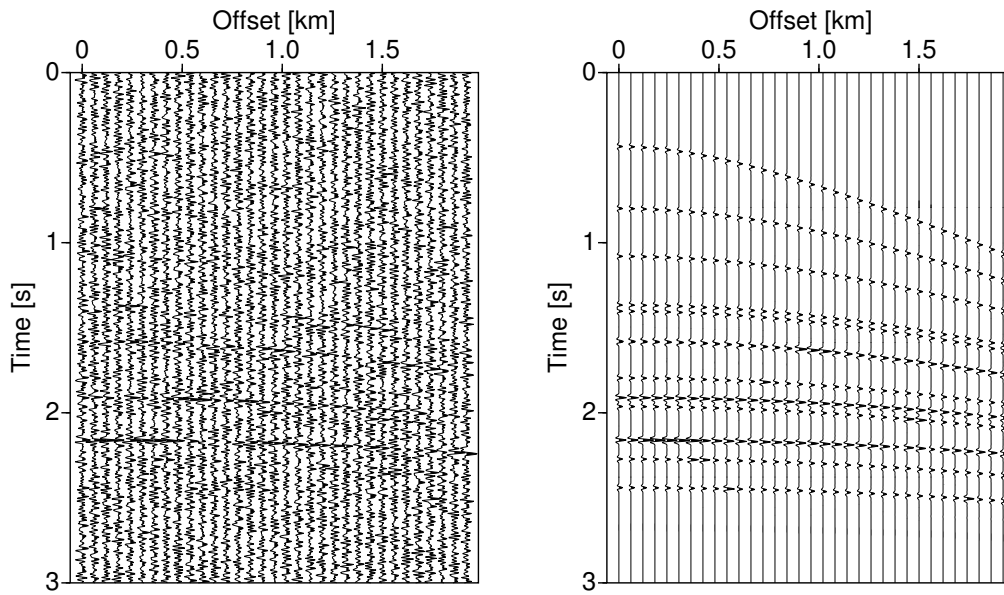


Figure 3.6: One selected CMP gather with and without random noise. The noise follows a uniform distribution and has been band-limited to the same frequency range as the prestack data without noise.

3.4.2 CRS stack results

The first step was the application of the CRS stack using the implementation described in Mann (2002). The search for the attribute triplet parameterising the best-fit CRS operator has been performed in two consecutive parts: a split search in subsets of the data for initial attribute values and a three parameter optimisation to get the final attribute values. These optimised attributes have been used to obtain the stacked section displayed in Figure 3.7. The coherence along the CRS stacking operator is shown in the same figure. Due to the very high noise level, the semblance values are relatively low. Coherent reflection events have been detected over the entire section and the signal-to-noise ratio in the final CRS stack section is highly increased compared to the prestack data.

The next step was the event-consistent smoothing of the stacking parameters using the robust line fit method described in Section 3.1.3. Output of this step were not only smoothed stacking parameter sections but also uncertainty estimates that were applied in the tomographic inversion. The smoothing was applied to the horizontal slowness of the normal ray and second traveltime derivatives with respect to the midpoint and half-offset coordinates. These derivatives are easily calculated from the kinematic wavefield attributes NIP-wave curvature radius, normal wave curvature radius, and normal ray emergence angle which are output of the used CRS stack implementation. The corresponding relationships are found in Section 3.1. The smoothing of and the uncertainty estimation for traveltime derivatives and slowness is more convenient as these quantities constitute the actual input for the tomographic inversion.

The sections of horizontal slowness p and the second traveltime derivative M before and after smoothing as well as the corresponding error estimate sections are shown in Figure 3.8. All samples for which no linear regression was performed are blanked out for a better visualisation. The improvements in attribute continuity due to the applied smoothing strategy can be clearly observed in the section of the second traveltime derivative M . Both, the error estimate for p and M , show strong lateral variability along the reflection events. Thus, they clearly demonstrate the influence of the noise on the

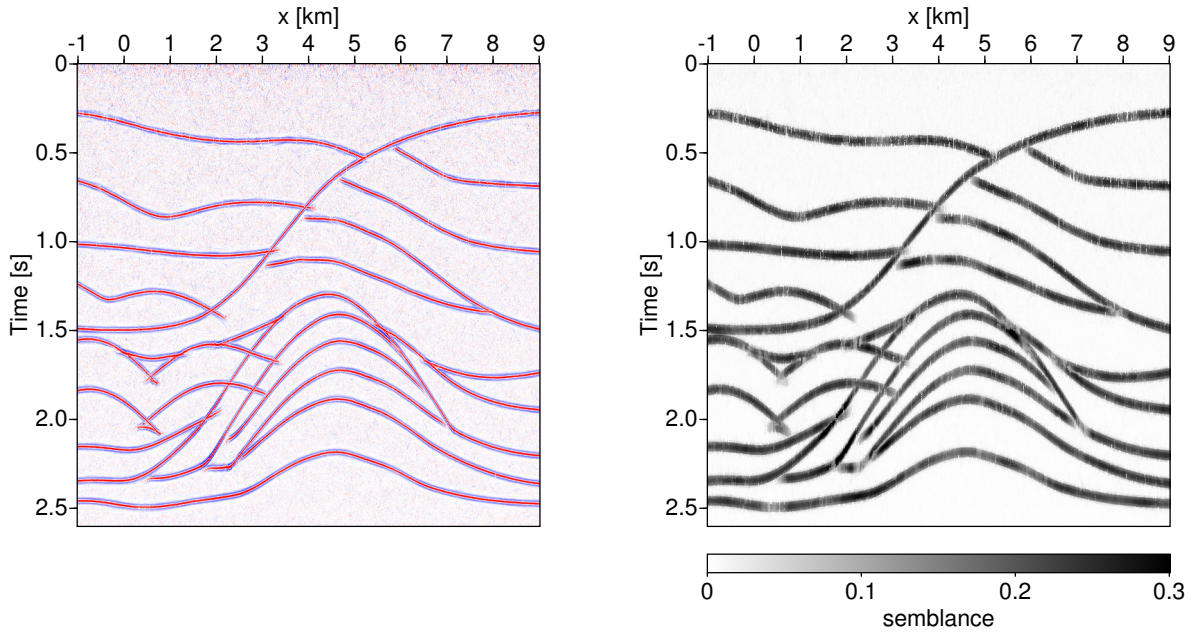


Figure 3.7: CRS stack section and associated coherence after simultaneous three-parameter optimisation.

attribute quality. A subset of the lowermost reflection event has been used for the demonstration of the event-consistent smoothing using local robust linear regression described in Section 3.1.3 and is displayed in Figure 3.4.

Whereas the error of the slowness p is in the same range in the entire section, the error of M shows a clear trend: it decreases with increasing traveltimes. This indicates that a small error in M at a large traveltimes changes the CRS operator in terms of moveout at a specific offset by the same amount as a large error in M at a small traveltimes. Thus, using these errors as weights in the inversion implicitly means to weight the uncertainty in operator moveout. Using constant weights for M in the inversion allows larger moveout changes of the CRS operator at large traveltimes. This results in a weighting of deep reflection events which is possibly too low.

The next step in the work flow for velocity model determination is the selection of reliable, at least locally coherent reflection events and the extraction of the associated information describing NIP-waves from the CRS attribute sections. For this purpose, I used the automated strategy described in Section 3.2.5. Together with the attributes themselves their uncertainty has been extracted. No further pick editing was necessary in this synthetic data example. Picking was performed only on every 15th trace of the CRS stack section. A total number of 455 reliable pick locations was detected.

These 455 picks formed the input for the tomographic inversion. The horizontal slowness p and second traveltimes derivative M associated with each pick have been weighted with the uncertainty estimated by the event-consistent smoothing strategy. The uncertainties for one-way traveltimes T and pick location x_m have been set to 0.002 s and 7.5 m, respectively. This corresponds to a picking error of one sample in the time dimension and of the half midpoint distance in the lateral dimension. The data classes have been balanced against each other by weighting them with the mean error of all their members. This was necessary to keep all numbers in the inversion algorithm in a comparable order of magnitude.

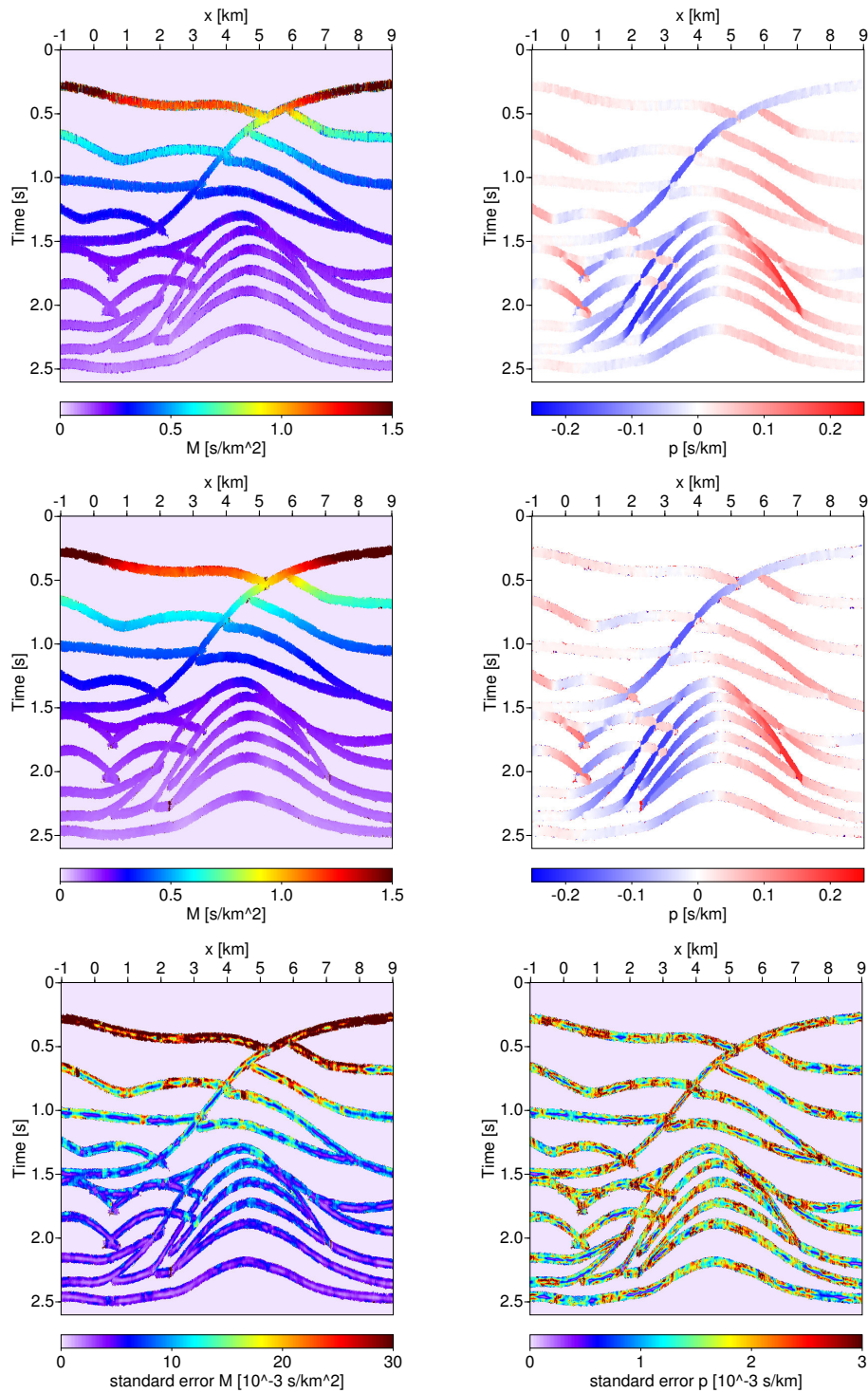


Figure 3.8: CRS stacking parameters M (left-hand column) and p (right-hand column) before (top row) and after (middle) application of the event-consistent smoothing using the local robust linear regression technique. The mean absolute deviation of the attributes is displayed in the bottom row. Samples for which the smoothing was not applied have been blanked out.

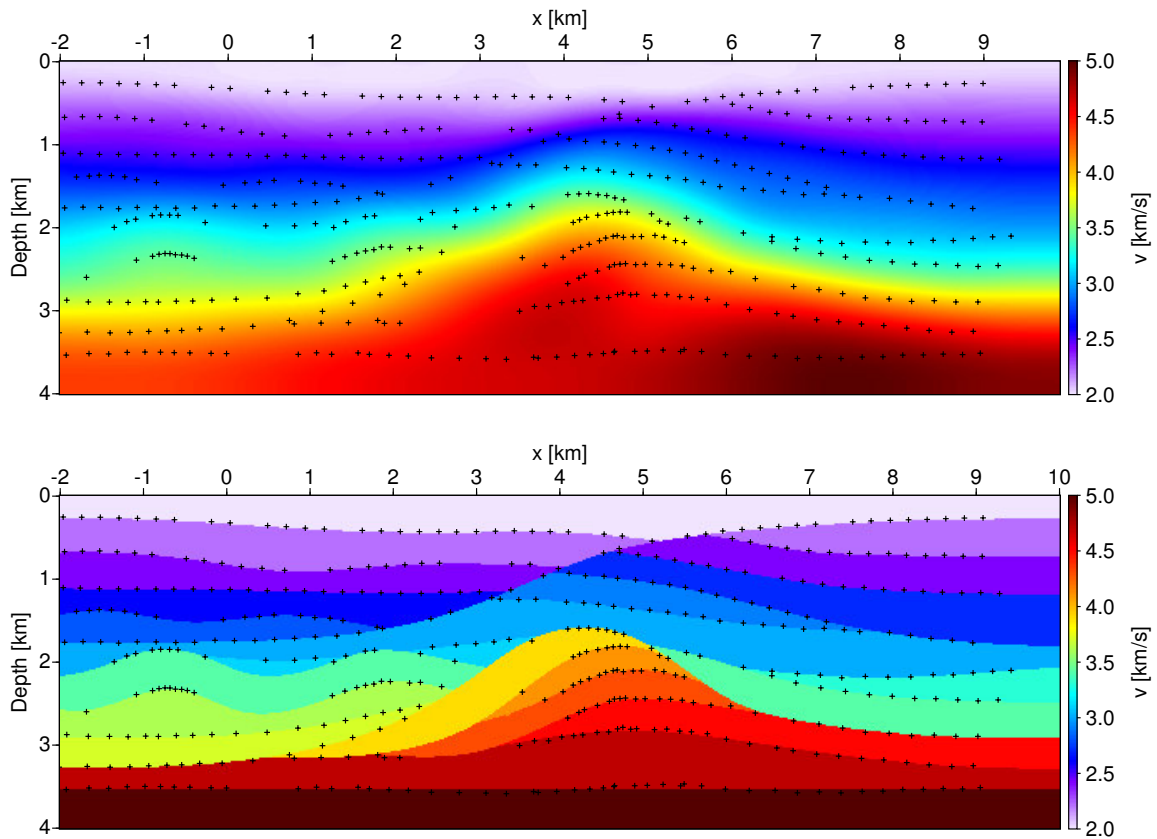


Figure 3.9: The final smooth velocity model obtained after 12 iterations (top) and the original blocky velocity model (bottom). On top of both, the obtained NIP locations are displayed.

The B-spline node grid of the velocity model description is built of $n_x = 21$ nodes in the horizontal direction and $n_z = 19$ nodes in the vertical direction. The horizontal nodes cover the range between -2100 m and 9100 m with a spacing of 400 m. The vertical nodes start at $z = 0$ m and end at $z = 3600$ m, the spacing between the nodes is 200 m. The B-spline coefficients were initialised using a vertical gradient model with $v = 2000$ m/s at the surface and a constant gradient of 0.67 s $^{-1}$.

The initial regularisation weights were set to $\epsilon_{xx} = 0.1$, $\epsilon_{zz} = 0.1$ and $\epsilon = 10^{-12}$. The weights for the minimisation of the second lateral derivatives of the velocity model have been decreased from iteration to iteration according to the square root of the fractional decrease in the cost-function value. The surface velocity has been constrained to achieve the true value of 2000 m/s. The corresponding weight was set to one. After twelve iterations the inversion has been stopped as the cost-function value did not further decrease.

The final smooth velocity distribution is shown in Figure 3.9 with the obtained NIP locations displayed on top. Comparing it with the original blocky velocity model shown in the same figure one can see that the overall reflector structure has been well recovered. The structure in the smooth velocity model is similar to the one in the true model. In the deep right-hand part of the inverted velocity model an artificial structure is observed which, however, does not prevent the NIP locations to follow the reflector structure. Together with the final residual error of the forward-modelled attributes which is much smaller than the a priori assumed one, these results demonstrate that it is possible to invert for depth velocity models using kinematic information describing NIP waves and the associated

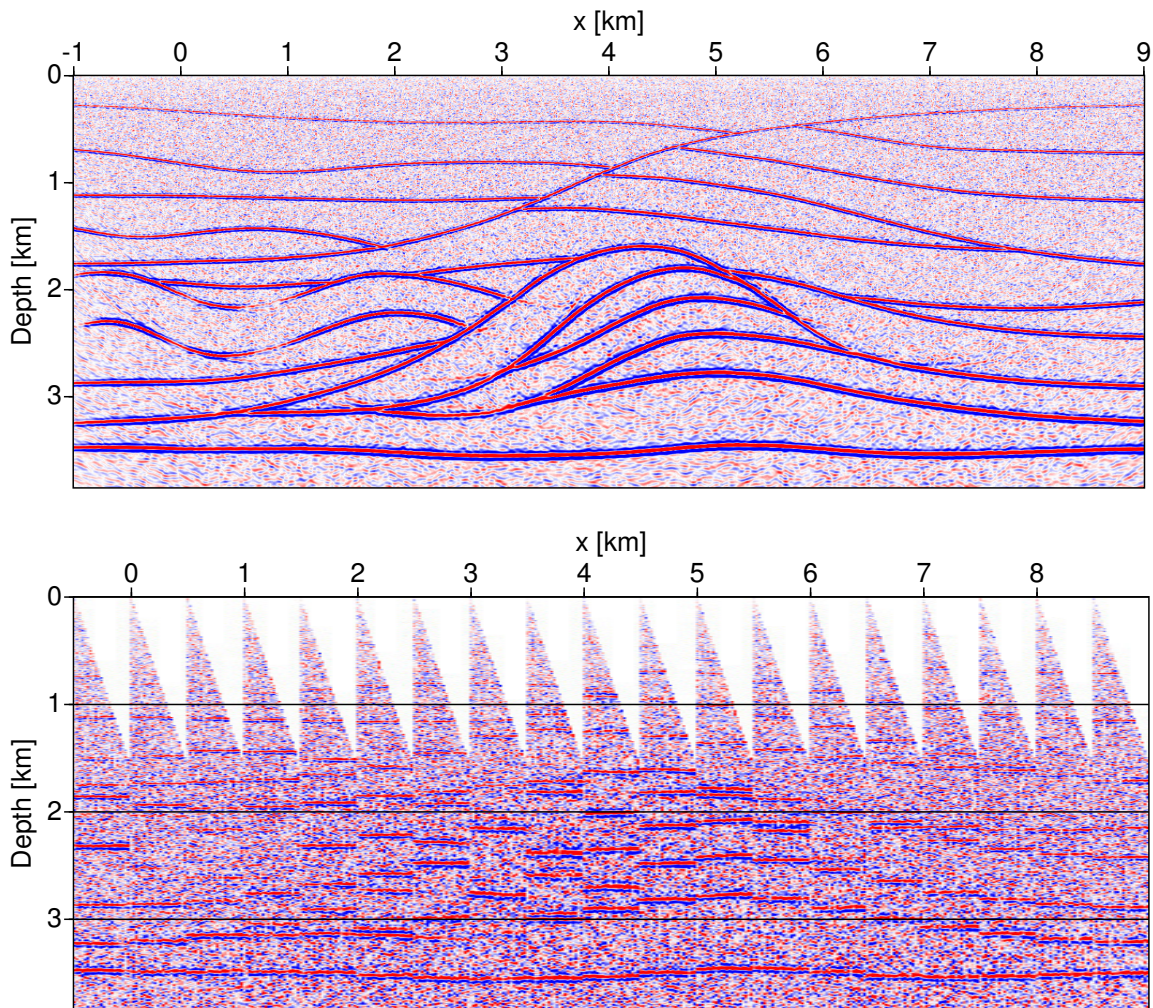


Figure 3.10: Stack of prestack depth migrated CO sections (top) and selected CIGs (bottom). The maximum offset displayed and used for migration is 2000 m.

uncertainty estimates.

The final check for the consistency of the velocity model with the seismic data is prestack depth migration. Offsets up to 2000 m have been used in the migration. Figure 3.10 shows CIGs with a spacing of 500 m and a stack of the prestack depth migrated common-offset sections. The CIGs are flat and, thus, indicate that the velocity model is consistent with the prestack data for the considered offset range. A comparison between the stack of the depth migrated common-offset sections and the true velocity model reveals that the overall reflector structure has been well recovered.

If migration is performed including larger offsets significant residual moveout would be detected in the CIGs. This is due to the fact that the method presented here is based on a second-order approximation of the diffraction traveltimes and that the quantities parametrising these traveltimes suffer from random noise and systematic effects like the spread-length bias. To overcome these limitations, a velocity model building technique using model-based diffraction traveltimes is presented in the next chapter which is suitable to update the models determined with NIP-wave tomography.

Chapter 4

Velocity model building from exact diffraction traveltimes

A new velocity model building technique is presented in the following which is based on forward-modelled diffraction traveltime functions. The method overcomes the limitation to second-order inherent in NIP-wave tomography (see previous chapter) as the model-based diffraction traveltime operators may be arbitrarily complex. The forward-modelled diffraction traveltime functions are used for a moveout correction of the prestack data. Criteria are defined in this moveout-corrected time domain which are used to minimise the kinematic difference between CRP gathers and CIGs. The tomographic inversion scheme presented here aims at enforcing the kinematic equivalence between the CRP gathers and CIGs for all considered reflection points which may be sparsely distributed.

The same model description as for NIP-wave tomography is used for the implementation. Each reflection point is treated independently of all others. There exists a mutual relationship between the analysis locations in depth and the pick locations in the ZO time domain which have been used in the formulation of NIP-wave tomography. This relationship is established by the normal ray. Thus, the here presented method is an ideal updating technique to further improve the velocity models obtained by the inversion scheme based on approximate diffraction traveltime functions.

After a motivation of the here presented approach, I extensively discuss the concept of using model-based diffraction traveltime operators as moveout functions. The result of this moveout correction is more easily interpreted in special coordinates which are unique for each considered depth point: the illumination and scattering angle. The whole method is formulated in the angle domain. The inverse problem to be solved is stated and data and model components of the inversion scheme are introduced. The chapter ends with the presentation of a 2D synthetic data example which demonstrates the applicability of the proposed method.

4.1 Motivation

Several common methods for the determination of velocity models suitable for depth migration are based on the analysis of the flatness of common image gathers (CIG). See Chapter 1 for a short overview of this subject. Flatness analysis is performed either for selected analysis locations or horizon-based in a layer stripping manner. The residual moveout in the CIGs, which is a function

of depth, may be converted to traveltime residuals using the local geological dip and velocity. These traveltime misfits then serve as input for a tomographic update of the velocity model (see, e. g., [Stork, 1992](#)).

This approach makes a fundamental assumption: a CIG contains all information stemming from the same depth point. This perfectly holds only for the case that the true velocity model is known. However, as this model generally can not be determined from the seismic data alone a velocity model is searched for which is as consistent as possible with the acquired seismic data and any a-priori knowledge available. As long as flatness analysis for CIGs is only performed at sparsely selected analysis locations, a CIG usually deviates kinematically from a common reflection point (CRP) gather. The CRP gather is composed of all traces containing signals from a selected reflection point in a specified depth model. The corresponding source and receiver locations may be determined by ray tracing in the depth model which consists of the velocity itself, the reflection point of interest, and the geological dip at that position. Thus, the CRP gather is model dependent. If the depth model is consistent with the prestack data for a selected reflection point, the CRP gather and the CIG for that position must be kinematically identical.

This condition ensures that all traces in the CRP gather contain a stationary point, i. e., a point of tangency between the model-based diffraction traveltime operator and the reflection event in the prestack data. The kinematic equivalence between CRP gather and CIG not only implies a flat CIG but also the same geological dip in all depth migrated common-offset or common-scattering-angle sections. This is a necessary condition for accurate imaging with depth migration.

4.2 Diffraction traveltimes as moveout function

The core of the new inversion scheme is the application of forward-modelled diffraction traveltime operators as moveout functions. In Kirchhoff depth migration ([Schneider, 1978](#)) the weighted sum of all signals spread along a diffraction traveltime operator is assigned to the corresponding depth position. If there exists a point of tangency between the diffraction traveltime operator and an event in the prestack data, i. e., a stationary point, this summation is constructive. A ZO stationary point is characterised by the ZO two-way traveltime, the trace position, and the local time dip of the reflection event in the ZO section. The corresponding depth position and geological dip may be found by normal ray tracing. Note that this relationship between the ZO time domain and the depth domain holds in principal for any velocity model. This means that for ZO the kinematic equivalence between a CRP gather and the corresponding CIG may be assured even for velocity models which are not consistent with the prestack data.

The relationship between time and depth domain for the ZO case is depicted in [Figure 4.1](#). Point N on the locally coherent event in the time domain (solid line in the upper half-space) is described by its lateral position x_{m_0} , the ZO two-way time $t_0 = 2 T_0$ and the local time dip $2 p_0$ of the event, where p_0 denotes the horizontal component of the slowness vector and T_0 is the ZO one-way traveltime. Associated with point N is the reflection point M in depth where the ZO ray (indicated by the dashed line connecting x_{m_0} and M) hits the reflector normally. The dashed half-circle in the depth domain indicates the ZO isochron, i. e., the line of equal two-way traveltime t_0 which is tangent to the reflector at point M. The counterpart to the isochron is the ZO diffraction traveltime curve depicted as dashed curved line in the time domain. The diffraction traveltime curve is tangent to the locally coherent event at the stationary point N. The local geological dip at reflection point M is given by Θ_0 . It is

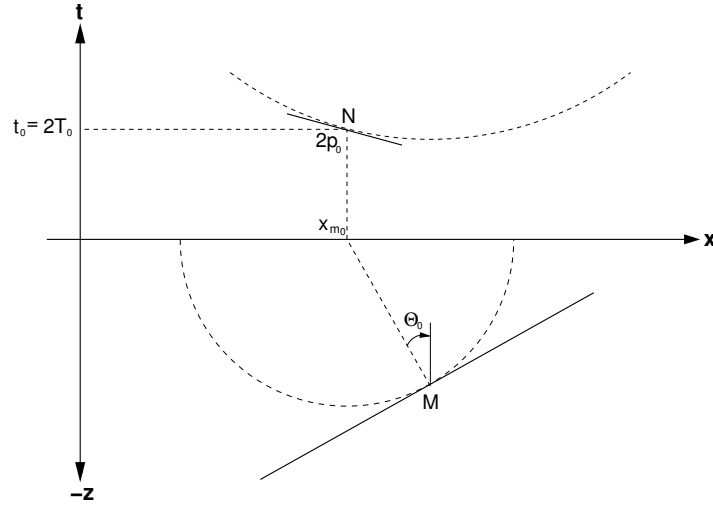


Figure 4.1: Relationship between time and depth domain for the ZO case. For visualisation purposes the depth domain is depicted with the negative z -axis pointing downwards. See main text for details.

measured from the normal ray to the positive vertical axis. This definition is used here for all rays starting at a selected depth point. The angle between an arbitrary ray and the vertical will be denoted by γ . Angles are defined to be positive if they are measured counterclockwise and vice versa. This definition ensures that the sign of the initial horizontal slowness component of a ray starting at point M equals the sign of the angle between the ray and the vertical, i. e., $p_x(M) = \sin(\gamma)/v(M)$, where $v(M)$ is the wave propagation velocity at point M .

Once the depth position corresponding to a selected pick in the ZO time domain has been found, the diffraction traveltimes for that position can be forward modelled, e. g., using ray tracing as it is presented in Chapter 2. The diffraction traveltime function τ_M for point M is given by traveltimes along all ray pairs connecting a source S with a receiver G , both at the surface, through point M . This means that the diffraction traveltimes are functions of source and receiver coordinates, i. e., $\tau_M = \tau_M(x_S, x_G)$. These, in turn, are functions of the corresponding ray take-off angles γ_S and γ_G at point M . Defining the scattering angle Φ and the illumination angle Θ by

$$\Phi = (\gamma_S - \gamma_G)/2 \quad \text{and} \quad \Theta = (\gamma_S + \gamma_G)/2, \quad (4.1)$$

the diffraction traveltime operator becomes a function of these angles: $\tau_M = \tau_M(\Theta, \Phi)$. For a graphical explanation of Θ and Φ see Figure 4.2. Note the analogy between the angle domain and the commonly used surface coordinates: shot and receiver coordinates at the surface correspond to take-off angles in depth, the midpoint coordinate $x_m = (x_S + x_G)/2$ corresponds to the illumination angle Θ , and the half-offset $h = (x_S - x_G)/2$ corresponds to the scattering angle Φ . The relationships between these quantities are non-linear functions of the subsurface velocity model. They are described in detail below.

The diffraction traveltime function τ_M is defined as the surface of traveltime zero for its application as moveout function. In this way, every trace with source and receiver locations within a considered range gets a different delay time. Traveltime is measured with respect to the diffraction traveltime operator. This means that the time axis directly displays differences between the diffraction traveltimes and the original time axis. This transformation is unique for a selected depth position and velocity model. It changes as soon as changes in the depth position and velocity model occur.

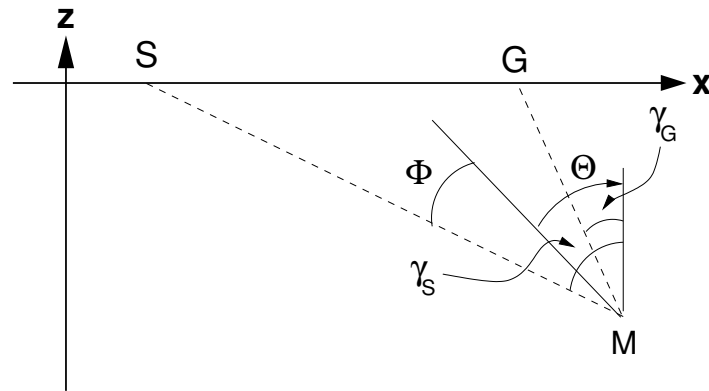


Figure 4.2: Definition of the illumination angle Θ and the scattering angle Φ . See main text for details.

4.2.1 Analytic example

I demonstrate this transformation using a simple model in which all quantities can be calculated analytically: point M is located at $x = 0$ m and $z = 3500$ m on a reflector with a dip of 15° and a homogeneous overburden with a wave propagation velocity of 2000 m/s. Firstly, I present the transformation results for the correct velocity model. Then, the corresponding results for a wrong velocity model are shown.

The traveltimes along specular ray pairs associated with the reflection event in the prestack data are visualised with isochrones as a function of midpoint and offset in Figure 4.3(a). In Figure 4.3(b) these traveltimes are shown as a function of scattering and illumination angle. Whereas the traveltimes correspond to specular reflections occurring anywhere on the reflector the angles correspond to rays which all start at point M and, thus, mostly are non-specular. These are the angles of the ray pairs building the diffraction traveltimes operator of point M. The isochrones along this operator are shown in Figure 4.3(c) in the midpoint-offset domain and in Figure 4.3(d) in the angle domain. Figure 4.3(e) and 4.3(f) show the difference between reflection and diffraction traveltimes in the midpoint-offset and angle domain, respectively. In all figures, the CRP trajectory is highlighted as dashed-dotted line. The CRP trajectory is given as that line along which the traveltime difference vanishes. Generally, it is a curved line in the midpoint-offset domain. In the angle domain it simplifies to a straight line because the illumination angle must equal the geological dip. This can clearly be seen in Figure 4.3(f).

Figure 4.4 shows time dips, i. e., the first partial derivative of traveltime with respect to the midpoint coordinate and the illumination angle, respectively. Sub-figures are arranged in the same way as in Figure 4.3. As the CRP trajectory is built of specular reflections, not only the traveltime difference but also the time dip difference vanishes along it as can be seen in Figure 4.4(e) and Figure 4.4(f). Vanishing time and time dip difference between reflection and diffraction traveltime surface along the CRP trajectory is a necessary condition for correct subsurface imaging. This simply states that the CRP trajectory is a trajectory of stationary points in a velocity model which is consistent with the data. Note that the difference between time dip with respect to offset and scattering angle automatically vanishes along the CRP trajectory as soon as there is no difference between reflection and diffraction traveltimes.

Figures 4.5 and 4.6 show the same quantities for a wrong depth model. The velocity has been changed to 2300 m/s. Point M has been chosen such that the ZO stationary point remains unchanged. There-

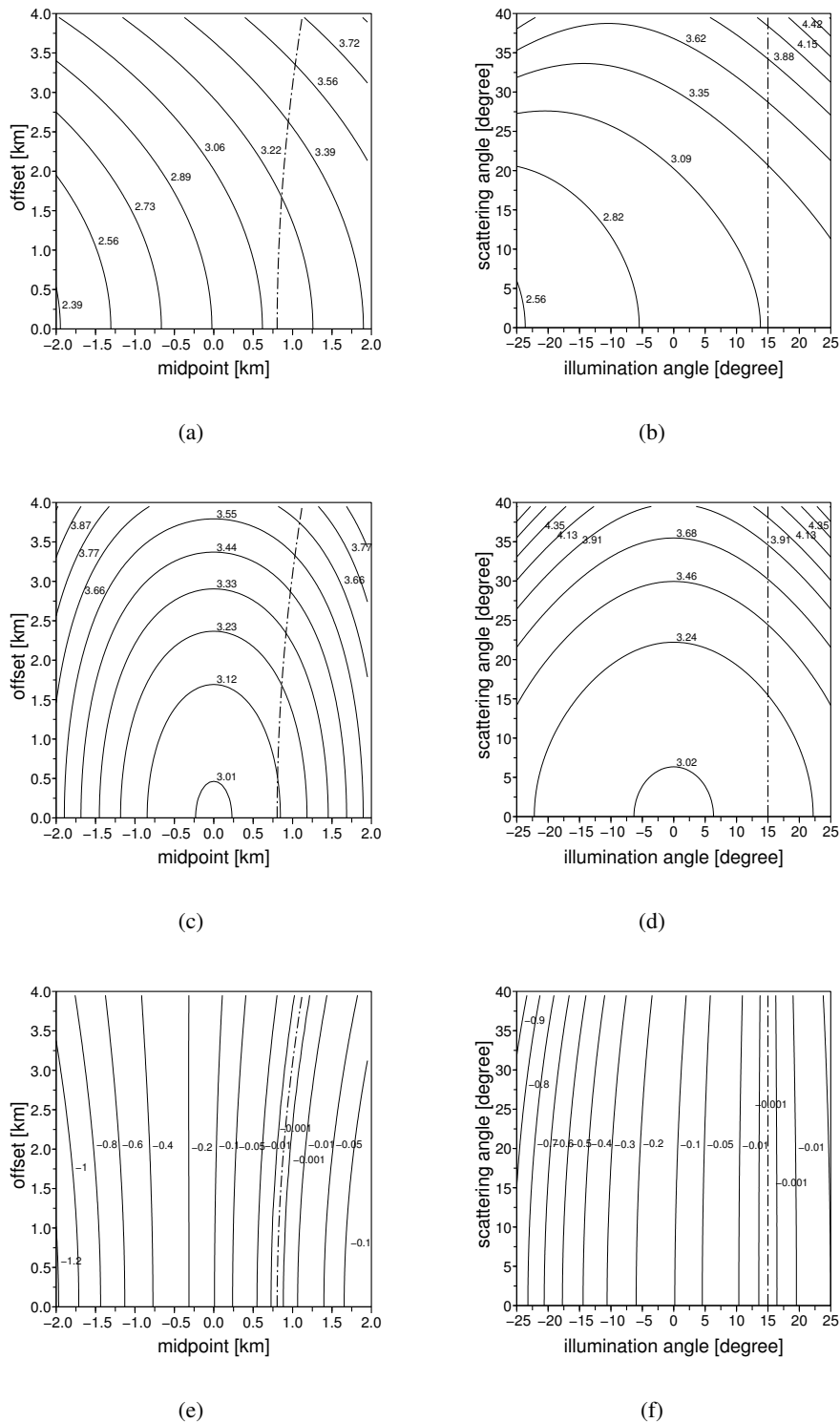


Figure 4.3: Correct velocity model: traveltime distribution [s] along the reflection event (a) in the midpoint-offset domain and (b) in the angle domain. Traveltime distribution along the diffraction traveltime operator (c) in the midpoint-offset domain and (d) in the angle domain. Traveltime difference between reflection event and diffraction traveltime operator (e) in the midpoint-offset domain and (f) in the angle domain. The CRP trajectory is indicated by the dashed line.

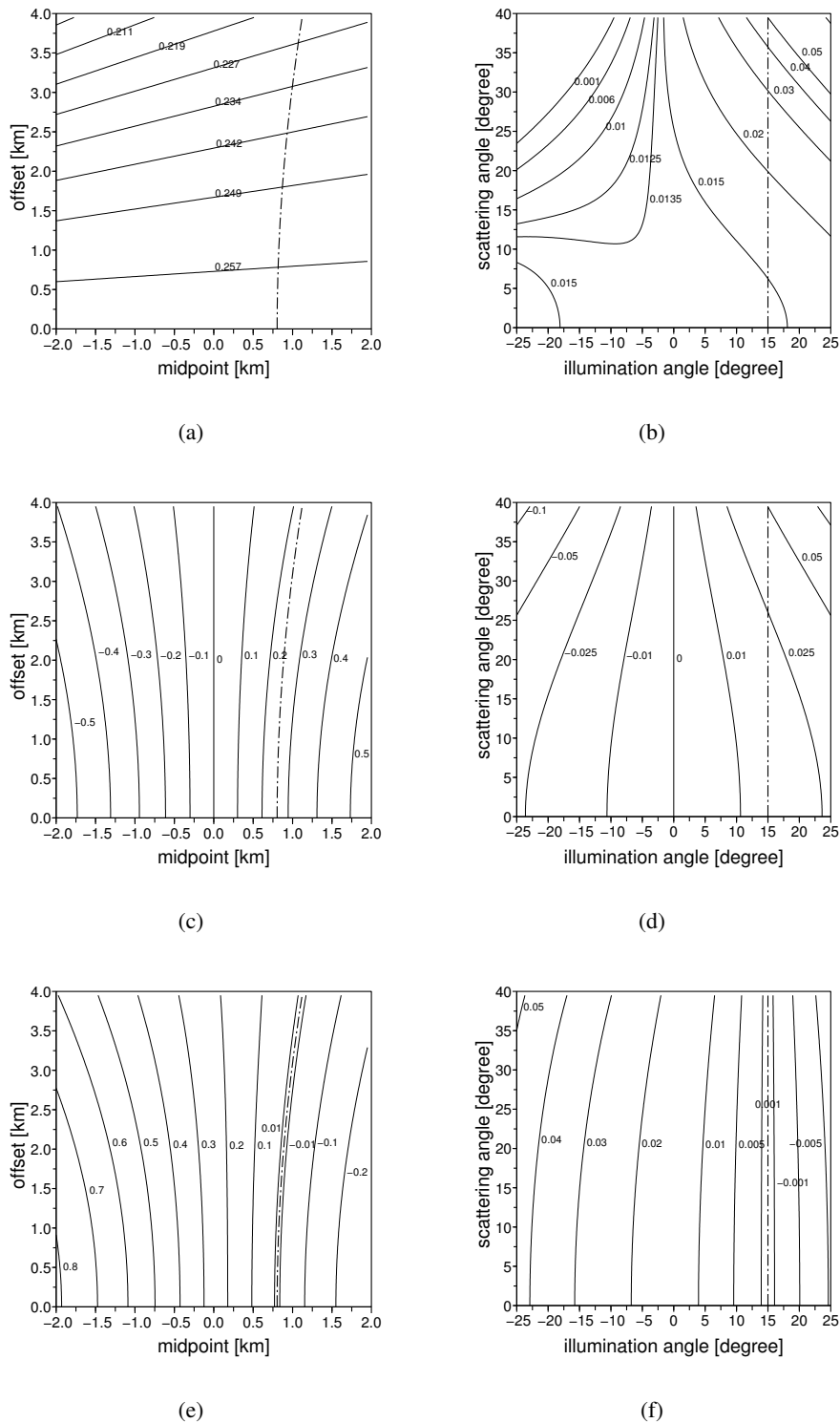


Figure 4.4: Correct velocity model: time dip distribution along the reflection event (a) in the midpoint-offset domain [s/km] and (b) in the angle domain [s/°]. Time dip distribution along the diffraction traveltime operator (c) in the midpoint-offset domain and (d) in the angle domain. Time dip difference between the reflection event and the diffraction traveltime operator (e) in the midpoint-offset domain and (f) in the angle domain. The CRP trajectory is indicated by the dashed line.

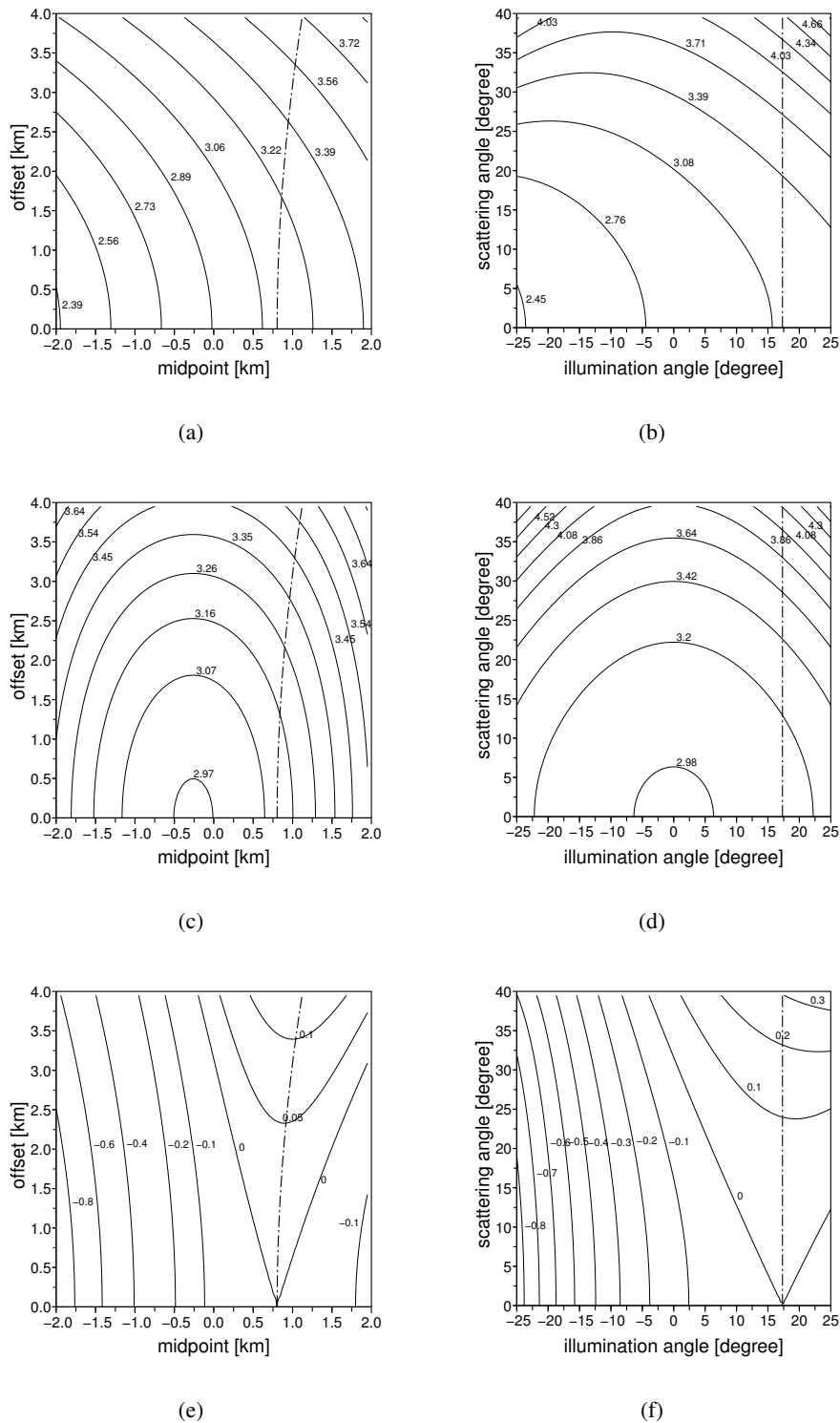


Figure 4.5: Wrong velocity model: traveltime distribution [s] along the reflection event (a) in the midpoint-offset domain and (b) in the angle domain. Traveltime distribution along the diffraction traveltime operator (c) in the midpoint-offset domain and (d) in the angle domain. Traveltime difference between reflection event and diffraction traveltime operator (e) in the midpoint-offset domain and (f) in the angle domain. The CRP trajectory is indicated by the dashed line.

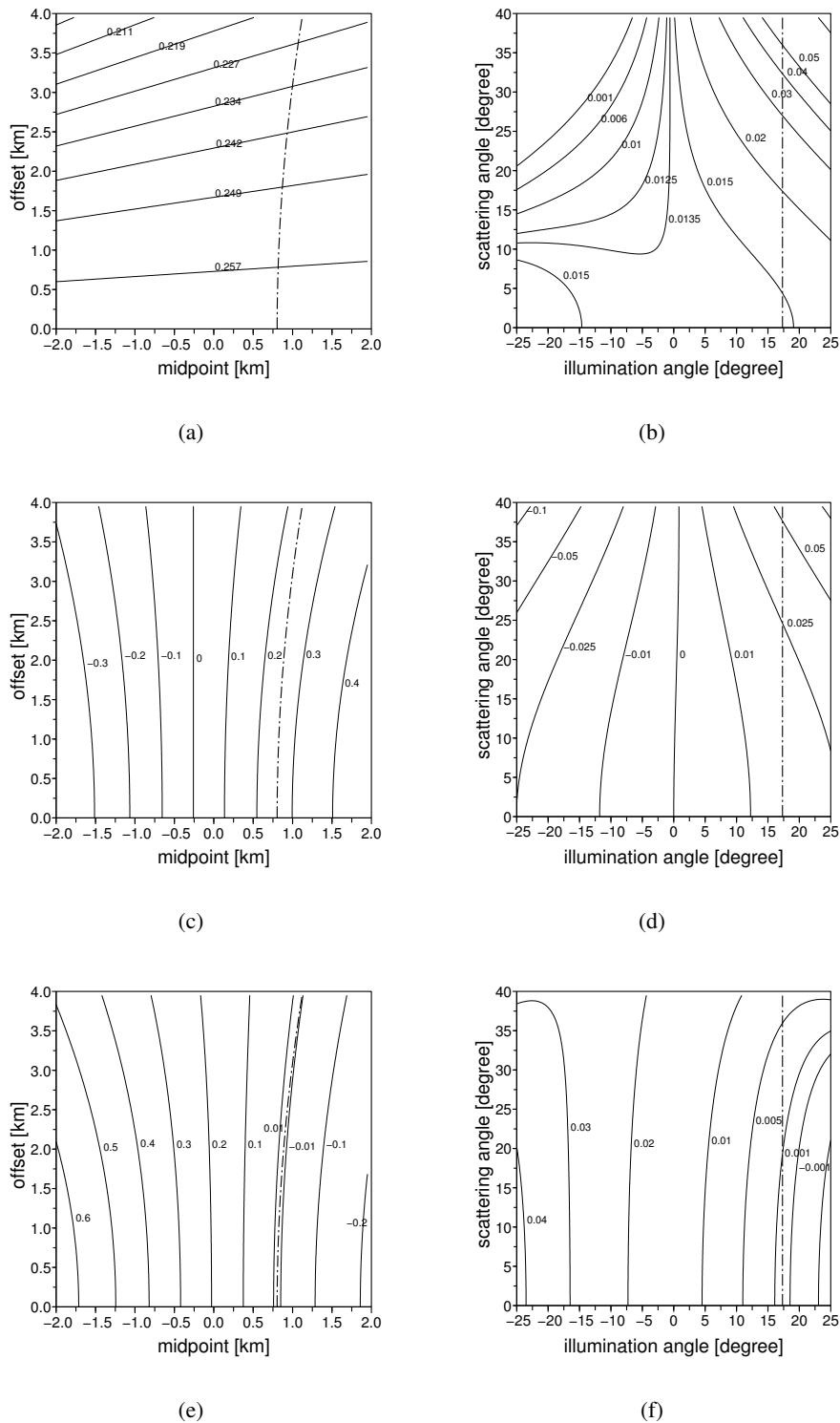


Figure 4.6: Wrong velocity model: time dip distribution along the reflection event (a) in the midpoint-offset domain [s/km] and (b) in the angle domain [$^{\circ}$]. Time dip distribution along the diffraction traveltimes operator (c) in the midpoint-offset domain and (d) in the angle domain. Time dip difference between the reflection event and the diffraction traveltimes operator (e) in the midpoint-offset domain and (f) in the angle domain. The CRP trajectory is indicated by the dashed line.

fore, the CIG and CRP gather are kinematically equivalent for ZO. However, this no longer holds for finite offsets and finite scattering angles. In a wrong depth model the time and time dip difference no longer vanishes along the CRP trajectory. This is demonstrated in Figures 4.5(e) and 4.5(f) for the time difference and in Figures 4.6(e) and 4.6(f) for the time dip difference. The reflection traveltimes are fixed as they depend on the true velocity model. The diffraction traveltimes, the illumination and scattering angles, and the CRP trajectory are model dependent quantities. Thus, they have changed in the wrong model. This leads to the deviation of the CRP trajectory from a trajectory of stationary points.

This simple data example clearly demonstrates the conditions necessary for the correct depth imaging of seismic data related to a single depth point and leads to the formulation of the tomographic inversion scheme.

4.3 The tomographic inversion scheme

The specular ray pairs for a reflection point M corresponding to a ZO stationary point are constrained to $\Theta = \Theta_0$. The traveltimes t along these specular ray pairs are given by the subset $t = \tau_M(\Theta = \Theta_0, \Phi)$ of the diffraction traveltime function. Source and receiver positions together with these traveltimes define the CRP trajectory along which the reflected energy stemming from point M can be found in the prestack data.

For the migration of this energy to reflection point M it is not sufficient to find a velocity model, depth position, and geological dip which only explains the CRP trajectory. It is also necessary that in each point of the CRP trajectory the diffraction traveltime function and the reflection event in the prestack data have the same time dip with respect to the midpoint or illumination angle, respectively. This is a necessary condition for a constructive summation of the energy stemming from reflection point M in the framework of Kirchhoff depth migration. It further implies that all stationary points are aligned along the CRP trajectory.

The goal of the inversion scheme presented in the following is to find a velocity model in which this kinematic equivalence of the CRP gather and the CIG is ensured for all selected picks in the ZO time domain. The quantities measuring the kinematic misfit between the CRP gather and the CIG are defined below. They are non-linear functions of the velocity model. Therefore, the inversion scheme proceeds iteratively by minimising a cost-function which measures the kinematic misfit between the CRP gather and the CIG. The cost-function is formulated in the least-squares sense.

4.3.1 Data and model components

The inversion starts with the selection of a number of picks in a ZO section which can be obtained by any kind of stacking procedure. For each pick all quantities describing the ZO stationary point have to be determined. In practice, I use the CRS stack method presented in Section 3.1 which automatically provides the needed quantities (Mann, 2002). Reliable events in the ZO stack section may be identified using the picking strategies described in Section 3.2.5. Initial estimates of the depth position and geological dip related to each pick in the ZO time domain are determined in the initial velocity model using ray tracing. The implementation of the method uses the same B-spline representation (de Boor, 1978) of the velocity model as used in NIP-wave tomography. The velocity

model description is treated in Appendix A. The identical velocity model parameterisation together with a similar description of the ZO picks makes the method presented in this chapter an ideal updating procedure for the velocity models based on approximate diffraction traveltime functions.

The complete model description consists of all reflection point positions, geological dips, and the B-spline coefficients representing the velocity distribution. All model components are summarised in the model vector \mathbf{m} :

$$\mathbf{m} = (x_i, z_i, \Theta_{0_i}, v_{jk})^T, \quad i = 1, \dots, n_{\text{data}}, \quad j = 1, \dots, n_x, \quad k = 1, \dots, n_z, \quad (4.2)$$

where n_{data} is the number of picks in the ZO section. The number of B-spline nodes in horizontal and vertical direction is given by n_x and n_z , respectively.

For each reflection point the diffraction traveltime operator is computed using ray tracing in the current velocity model. All traces within a user-defined aperture centred at each ZO pick location x_{m_0} are moveout corrected using the corresponding forward-modelled diffraction traveltime function and sorted with respect to the illumination and scattering angle. Only a small time window centred at the diffraction traveltimes is kept. In this way, a small time-shifted prestack data cube in the angle domain is created for each pick. Inside these cubes time and time dip residuals for all scattering angles are extracted along the model-based CRP trajectory. The necessary picking strategy is described below. Generally, this leads to a different number of picks in each moveout corrected cube. All time and time dip residuals are summarised in a data misfit vector:

$$\Delta \mathbf{d}_{\text{Cube}} = \left(\Delta t_i(\Theta_{0_i}, \Phi_l), \frac{\partial}{\partial \Theta} \Delta t_i(\Theta_{0_i}, \Phi_l) \right)^T, \quad i = 1, \dots, n_{\text{data}}, \quad l = 1, \dots, n_{\text{pick}_i}, \quad (4.3)$$

where n_{pick_i} denotes the number of picks in the i -th moveout corrected prestack data cube. These data misfits are used to update the depth model. The update is performed iteratively due to the non-linearity of the inverse problem. Once the model has been updated new estimates of the depth positions and geological dips may be obtained from the corresponding ZO picks by normal ray tracing. Alternatively, the estimates for the reflection point positions and geological dips obtained by the inversion can be used. In the latter case the inversion scheme allows for errors in the description of the ZO stationary points. As outlined in Section 3.1.3, the kinematic quantities describing a ZO stationary point suffer from errors due to, e. g., low signal-to-noise ratio in the seismic prestack data. Therefore, the misfit of the quantities describing the ZO stationary point

$$\Delta \mathbf{d}_{\text{ZO}} = (\Delta T_{0_i}, \Delta p_{0_i}, \Delta x_{m_{0_i}})^T, \quad i = 1, \dots, n_{\text{data}}, \quad (4.4)$$

is part of the complete data misfit vector given by

$$\Delta \mathbf{d} = (\Delta \mathbf{d}_{\text{ZO}}, \Delta \mathbf{d}_{\text{Cube}})^T. \quad (4.5)$$

Note that only the ZO data themselves are actually known. All finite offset quantities are unknown. Directly their misfit is picked in each iteration. The true finite-offset data do not need to be known for the inversion scheme. However, they can be calculated using the measured data misfit and the forward-modelled diffraction traveltime operator. The necessary equations are derived in Section 4.3.2. The ZO data misfit is easily determined using forward modelling by ray tracing.

With the model and data misfit vectors defined above, the inverse problem may be formulated as follows: find a model described by model vector (4.2) in which the misfit between the picked ZO

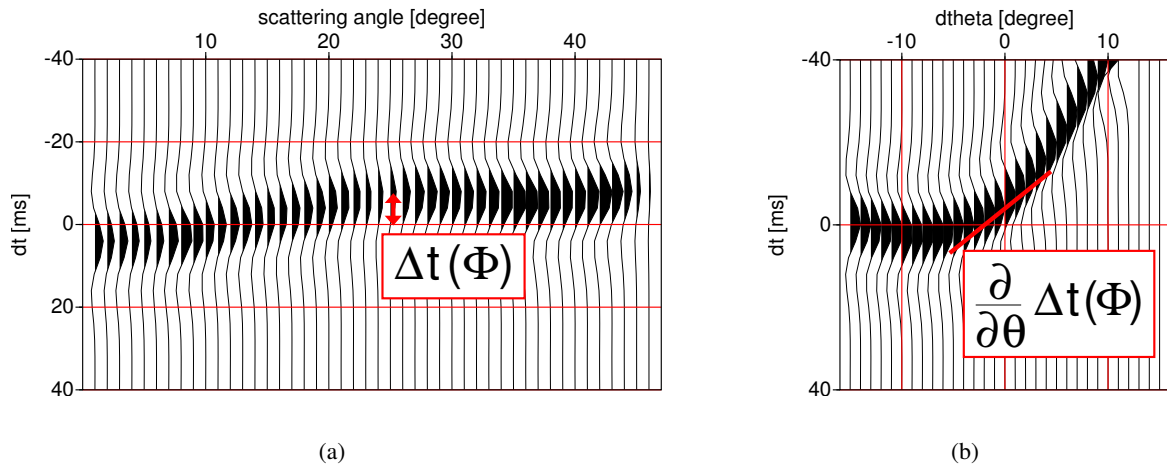


Figure 4.7: The (a) CRP gather and (b) a common-scattering-angle gather extracted from the moveout corrected prestack data cube. The moveout correction was performed with a velocity model inconsistent with the seismic prestack data. The illumination angle in (b) is measured with respect to the estimated geological dip, i. e., $d\theta = \Theta - \Theta_0$. Traveltime residuals $\Delta t(\Phi)$ are measured in the CRP gather whereas time dip residuals $\frac{\partial}{\partial \theta} \Delta t(\Phi)$ are measured in common-scattering-angle gathers. The difference between recording time and diffraction traveltimes is denoted by dt .

data components and their forward-modelled counterparts as well as the picked time and time dip residuals in all moveout corrected prestack data cubes is minimised. Like in NIP-wave tomography (see previous chapter) this problem is formulated as the minimisation of a cost-function in the least-squares sense. The solution as well as the regularisation and additional constraints that can be applied are described in Chapter 5.

4.3.2 Picking of time and time dip residuals

Residual moveout is mostly parameterised by an analytic function, e. g. a parabola. The parameters describing that function can then be determined by coherence analysis. However, when the residual moveout deviates from an analytic shape wrong results are obtained. Such a deviation is common in complex models and we observe it in Figure 4.7(a) which shows a CRP gather taken from the data example presented at the end of this chapter. For that reason, I propose a different picking strategy: each trace in the CRP gather is cross-correlated with a part of the ZO trace centred around the picked stationary point. I assume that the ZO trace provides a good estimate of the wavelet and that phase changes with increasing scattering angle of the wavelet in the CRP gather may be neglected. Picking the maximum in the cross-correlation panel directly gives the time residuals for all scattering angles. The cross-correlation is facilitated by the fact that the CRP gather shows no pulse stretch with increasing scattering angle as it is observed in a conventional CIG. This is due to the fact that the original sampling of the prestack data is preserved in the moveout correction with the model-based diffraction traveltime function.

Once the traveltime residuals have been extracted for all scattering angles, the corresponding time dip residuals are estimated by semblance maximisation (Neidell and Taner, 1971) along straight lines

centred around the a priori geological dip and the picked traveltime misfit in a common-scattering-angle gather. Such a gather is shown in Figure 4.7(b) together with the best fit line segment.

The proposed picking strategy works automatically. The results should be carefully examined to remove any outliers. This step requires a lot of human interaction.

4.3.3 The tomographic inversion algorithm

The data misfit vector (4.5) is used for the minimisation of cost-function Ψ defined in Chapter 5. In contrast to the used algorithm in NIP-wave tomography much more human interaction is required for the inversion scheme presented here. Due to the non-linearity of the inverse problem, the tomographic inversion algorithm proceeds iteratively:

- The B-spline representation of the velocity model is set up by defining the B-spline node sequence in each spatial direction and by assigning appropriate values to the B-spline coefficients. The complete B-spline representation may be copied from the result of NIP-wave tomography if the here presented method is used as an updating technique. The components of the data covariance matrix \mathbf{C}_D are defined. Additional constraints and appropriate weights for the regularisation have to be specified.
- Initial reflection point locations and geological dips are computed using ray tracing in the initial model: a normal ray, starting at the ZO surface location and in the opposite direction as defined by the corresponding ZO horizontal slowness, is traced downwards until the one-way time is depleted. The ray end point defines the initial reflection point location. The geological dip is calculated from the ray slowness vector. In this case, the initial misfit for the ZO data components vanish. When the here presented method is used as an updating technique initial estimates for the reflection point locations and geological dips as well as the misfit of the ZO data components are readily available from the results of NIP-wave tomography.
- The diffraction traveltime operators for all reflection points are forward-modelled using ray tracing. They are used for a moveout correction of all traces in the prestack data for which the source and receiver locations are within a user-defined aperture centred around the ZO location.
- Time and time dip residual misfits as a function of scattering angle are extracted along the CRP trajectory from the moveout corrected prestack data cubes using the automatic picking strategy described in Section 4.3.2. All picks are examined to remove outliers.
- The Fréchet derivatives for all data components are calculated using dynamic ray tracing and ray perturbation theory.
- The regularisation as well as additional constraints and their associated misfit are calculated. The cost-function Ψ is evaluated.
- The linear system (5.16) is solved using the LSQR algorithm (Paige and Saunders, 1982a,b). The obtained model update vector $\Delta \mathbf{m}$ is scaled with a factor $0 < \lambda \leq 1$ and added to the current model to obtain the new model to be used in the next iteration.
- New reflection point locations and geological dips are computed using ray tracing in the updated model. Alternatively, the corresponding estimates obtained by the inversion can be used.

- The diffraction traveltimes operators for the updated reflection point positions are forward modelled in the updated velocity model and applied as moveout functions to the prestack data.
- Time and time dip residual misfits are extracted along the CRP trajectory from the moveout corrected prestack data cubes. The misfit of the ZO data components is calculated using ray tracing in the updated model. The new cost-function value is estimated.
- If the cost-function does not decrease significantly and the flatness of the CRP gathers is not further improved the inversion is terminated. Otherwise, the Fréchet-derivatives are calculated and the next iteration is started with the inversion of the tomographic system (5.16). Note, that there is no automatic recalculation of the cost-function as this requires user interaction to examine the results of the automatic picking of time and time dip residual misfits.

4.3.4 Forward modelling and calculation of Fréchet derivatives

The forward modelling of diffraction traveltimes operators and ZO data components as well as the calculation of the Fréchet derivatives is based on ray and ray perturbation theory in ray-centred coordinates as it is described in Chapter 2.

Only the ZO data components actually need to be forward modelled. The traveltimes $t(\Theta_0, \Phi)$ and angle-domain time dips $\frac{\partial}{\partial \Theta} t(\Theta_0, \Phi)$ along the CRP trajectory do not need to be known for the tomographic inversion as directly their misfit is extracted from the moveout corrected prestack data cubes. The forward modelling of the ZO data components has already been outlined in Section 3.2.6 and is not described here, again.

The Fréchet derivatives are the first partial derivatives of all data components with respect to the model vector components. They may be symbolically written as

$$\frac{\partial \left(T_{0_i}, p_{0_i}, x_{m_{0_i}}, t_i(\Theta_{0_i}, \Phi_l), \frac{\partial}{\partial \Theta} t_i(\Theta_{0_i}, \Phi_l) \right)}{\partial (x, z, \Theta_0, v_{jk})}, \quad (4.6)$$

where $i = 1, \dots, n_{\text{data}}$, $l = 1, \dots, n_{\text{pick}_i}$, $j = 1, \dots, n_x$, and $k = 1, \dots, n_z$. For the calculation of the Fréchet derivatives a relationship between the angle-domain time dip and horizontal slowness components at the source and receiver locations is derived in Section 4.4. The partial derivatives of these slowness components with respect to the model components is directly available from dynamic ray tracing. The actual computation of the Fréchet derivatives is treated in Appendix B.

The partial derivatives of the ZO data components as well as the finite scattering angle traveltimes and time dips related to a selected reflection point are independent of the model components associated with all other reflection points. Thus, the tomographic matrix containing the Fréchet derivatives has a sparse structure. The implementation of the inversion takes this sparseness into account as only non-zero elements are actually kept in memory.

4.3.5 The structure of the tomographic matrix

The tomographic matrix has a similar structure as that in NIP-wave tomography. Using the short notations $t_{il} = t_i(\Theta_{0_i}, \Phi_l)$ and $t_{t,\Theta_{il}} = \partial t_{il} = t_i(\Theta_{0_i}, \Phi_l)/\partial \Theta$, the tomographic matrix reads:

$$\hat{\mathbf{F}} = \begin{pmatrix} w_T^{ZO} \frac{\partial T_i}{\partial x_j} \delta_{ij} & w_T^{ZO} \frac{\partial T_i}{\partial z_j} \delta_{ij} & w_T^{ZO} \frac{\partial T_i}{\partial \Theta_{0_j}} \delta_{ij} & w_T^{ZO} \frac{\partial T_i}{\partial v_{jk}} \\ w_p^{ZO} \frac{\partial p_i}{\partial x_j} \delta_{ij} & w_p^{ZO} \frac{\partial p_i}{\partial z_j} \delta_{ij} & w_p^{ZO} \frac{\partial p_i}{\partial \Theta_{0_j}} \delta_{ij} & w_p^{ZO} \frac{\partial p_i}{\partial v_{jk}} \\ w_x^{ZO} \frac{\partial x_{m_i}}{\partial x_j} \delta_{ij} & w_x^{ZO} \frac{\partial x_{m_i}}{\partial z_j} \delta_{ij} & w_x^{ZO} \frac{\partial x_{m_i}}{\partial \Theta_{0_j}} \delta_{ij} & w_x^{ZO} \frac{\partial x_{m_i}}{\partial v_{jk}} \\ w_T^{FO} \frac{\partial t_{il}}{\partial x_j} \delta_{ij} & w_T^{FO} \frac{\partial t_{il}}{\partial z_j} \delta_{ij} & w_T^{FO} \frac{\partial t_{il}}{\partial \Theta_{0_j}} \delta_{ij} & w_T^{FO} \frac{\partial t_{il}}{\partial v_{jk}} \\ w_{t,\Theta}^{FO} \frac{\partial t_{\Theta_{il}}}{\partial x_j} \delta_{ij} & w_{t,\Theta}^{FO} \frac{\partial t_{\Theta_{il}}}{\partial z_j} \delta_{ij} & w_{t,\Theta}^{FO} \frac{\partial t_{\Theta_{il}}}{\partial \Theta_{0_j}} \delta_{ij} & w_{t,\Theta}^{FO} \frac{\partial t_{\Theta_{il}}}{\partial v_{jk}} \\ & \mathbf{0}_v & & w_v \frac{\partial v^{\alpha-priori}}{\partial v_{jk}} \\ & \mathbf{0}_{v_q} & & w_{v_q} \frac{\partial |\nabla_{\perp} v|_i}{\partial v_{jk}} \\ & \mathbf{0}_B & & \mathbf{B} \end{pmatrix}, \quad (4.7)$$

with

$$\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

Each sub-matrix containing the partial derivatives of one class of the ZO data (like, e. g., traveltime T_0) with respect to the reflection point related part of the model description contains $n_{data} \times n_{data}$ elements. Each ZO data class has its own constant weight denoted by w_T^{ZO} , w_p^{ZO} , and w_x^{ZO} , respectively. Alternatively, a combination of data-driven uncertainty estimates with these constant weights can be used as it is applied in NIP-wave tomography. The sub-matrices containing the corresponding derivatives with respect to the B-spline coefficients each contain $n_{data} \times n_x n_z$ elements. The size of the part of the tomographic matrix related to the ZO data components is fixed except for the case that a ZO pick is removed between two iterations of the inversion. The sub-matrices related to the time and time dip misfits picked in the moveout corrected prestack data cubes may change their size from iteration to iteration. They depend on how many outliers are removed in the picking stage. Usually, their size increases from iteration to iteration as the velocity model becomes more consistent with the prestack data and the used scattering angle range can be increased. The size of these sub-matrices containing partial derivatives with respect to the reflection point related part of the model description is $n_{data} \times \sum_i^{n_{data}} n_{pick_i}$. The corresponding sub-matrices containing partial derivatives with respect to the B-spline coefficients have $n_x n_z \times \sum_i^{n_{data}} n_{pick_i}$ elements. Constant weights are applied to all time and time dip residuals misfits denoted by w_T^{FO} and $w_{t,\Theta}^{FO}$, respectively. The partial derivatives of the additional constraints (described in Chapter 5) with respect to the NIP-related model components vanish. The remaining sub-matrices containing the derivatives with respect to the B-spline coefficients have sizes of $n_{vdata} \times n_x n_z$ and $n_{vqdata} \times n_x n_z$, respectively. The remaining non-zero part is built by matrix \mathbf{B} related to the regularisation (see Chapter 5) and contains $n_x n_z \times n_x n_z$ elements.

4.4 Relationship between time dip residuals in the angle domain and in the midpoint-offset domain

The diffraction travelttime function $\tau(\Theta, \Phi)$ can be decomposed into source and receiver parts:

$$\tau(\Theta, \Phi) = \tau_S(\Theta, \Phi) + \tau_G(\Theta, \Phi). \quad (4.8)$$

Using the definition (4.1) the time dip of τ with respect to the illumination angle Θ reads

$$\frac{\partial \tau}{\partial \Theta} = \frac{\partial \tau_S}{\partial \Theta} + \frac{\partial \tau_G}{\partial \Theta} = \frac{\partial \tau_S}{\partial \gamma_S} \frac{\partial \gamma_S}{\partial \Theta} + \frac{\partial \tau_G}{\partial \gamma_G} \frac{\partial \gamma_G}{\partial \Theta} = \frac{\partial \tau_S}{\partial \gamma_S} + \frac{\partial \tau_G}{\partial \gamma_G}. \quad (4.9)$$

The emergence locations x_S and x_G of shot and receiver rays are functions of their take-off angle in depth. Therefore, equation (4.9) may be rewritten in a form relating time dip with respect to illumination angle Θ and horizontal slowness components at source and receiver locations:

$$\frac{\partial \tau}{\partial \Theta} = \frac{\partial \tau_S}{\partial x_S} \frac{\partial x_S}{\partial \gamma_S} + \frac{\partial \tau_G}{\partial x_G} \frac{\partial x_G}{\partial \gamma_G} = \pi_S \frac{\partial x_S}{\partial \gamma_S} + \pi_G \frac{\partial x_G}{\partial \gamma_G}, \quad (4.10)$$

where π_S and π_G denote the horizontal slowness of source and receiver ray building the diffraction travelttime operator. The same derivation can be performed for $\partial \tau / \partial \Phi$ yielding

$$\frac{\partial \tau}{\partial \Phi} = \frac{\partial \tau_S}{\partial x_S} \frac{\partial x_S}{\partial \gamma_S} - \frac{\partial \tau_G}{\partial x_G} \frac{\partial x_G}{\partial \gamma_G} = \pi_S \frac{\partial x_S}{\partial \gamma_S} - \pi_G \frac{\partial x_G}{\partial \gamma_G}. \quad (4.11)$$

Analogous equations are derived for the reflection travelttimes. However, as already stated above the angles correspond to non-specular ray pairs. Using the horizontal slowness components p_S and p_G the reflection time dip with respect to illumination angle Θ and scattering angle Φ reads

$$\frac{\partial t}{\partial \Theta} = p_S \frac{\partial x_S}{\partial \gamma_S} + p_G \frac{\partial x_G}{\partial \gamma_G} \quad \text{and} \quad \frac{\partial t}{\partial \Phi} = p_S \frac{\partial x_S}{\partial \gamma_S} - p_G \frac{\partial x_G}{\partial \gamma_G}. \quad (4.12)$$

The same analogy between quantities in the angle and midpoint-offset domain is observed that already has been seen in the definition of the illumination and the scattering angle: the time dip with respect to the illumination angle is the sum of the horizontal slowness components at the source and receiver locations, each slowness scaled by a model-dependent transformation factor. This is similar to the time dip with respect to the midpoint coordinate which is simply given as the sum of the horizontal slowness at the source and the receiver location. The same analogy is found between the time dip with respect to the scattering angle and the offset. The model-dependent factor is the transformation between surface coordinates and polar coordinates defined at the reflection point. It is easily computed using dynamic ray tracing (see Chapter 2).

The difference in time dip with respect to the illumination angle Θ between the reflection event and the diffraction travelttime operator may be written as

$$\begin{aligned} \frac{\partial}{\partial \Theta} \Delta t(\Theta, \Phi) &= \frac{\partial}{\partial \Theta} t(\Theta, \Phi) - \frac{\partial}{\partial \Theta} \tau(\Theta, \Phi) = \\ &= (p_S - \pi_S) \frac{\partial x_S}{\partial \gamma_S} + (p_G - \pi_G) \frac{\partial x_G}{\partial \gamma_G} = \Delta p_S \frac{\partial x_S}{\partial \gamma_S} + \Delta p_G \frac{\partial x_G}{\partial \gamma_G}. \end{aligned} \quad (4.13)$$

A similar expression is derived in the same way for the time dip difference with respect to the scattering angle Φ :

$$\begin{aligned} \frac{\partial}{\partial \Phi} \Delta t(\Theta, \Phi) &= \frac{\partial}{\partial \Phi} t(\Theta, \Phi) - \frac{\partial}{\partial \Phi} \tau(\Theta, \Phi) = \\ &= (p_S - \pi_S) \frac{\partial x_S}{\partial \gamma_S} - (p_G - \pi_G) \frac{\partial x_G}{\partial \gamma_G} = \Delta p_S \frac{\partial x_S}{\partial \gamma_S} - \Delta p_G \frac{\partial x_G}{\partial \gamma_G}. \end{aligned} \quad (4.14)$$

To obtain these time dip differences along the CRP trajectory, all derivatives in equation (4.13) and equation (4.14) have to be taken at the geological dip $\Theta = \Theta_0$.

For the tomographic inversion scheme presented here, only the determination of differences in the time dip with respect to the illumination angle Θ along the CRP trajectory is necessary. However, when the time dip differences with respect to the scattering angle Φ are known as well, the horizontal slowness components p_S and p_G can be determined along the CRP trajectory from the measured time dip residuals. Taking the sum of equations (4.13) and (4.14) yields twice the horizontal slowness error at the source scaled by the model-based coordinate transformation factor between surface coordinates and polar coordinates defined at the reflection point:

$$\frac{\partial}{\partial \Theta} \Delta t(\Theta, \Phi) + \frac{\partial}{\partial \Phi} \Delta t(\Theta, \Phi) = 2 \Delta p_S \frac{\partial x_S}{\partial \gamma_S}. \quad (4.15)$$

An analogous expression is derived for the horizontal slowness misfit at the receiver taking the difference of equations (4.13) and (4.14):

$$\frac{\partial}{\partial \Theta} \Delta t(\Theta, \Phi) - \frac{\partial}{\partial \Phi} \Delta t(\Theta, \Phi) = 2 \Delta p_G \frac{\partial x_G}{\partial \gamma_G}. \quad (4.16)$$

Along with the horizontal slowness components π_S and π_G related to the diffraction event and the model-based coordinate transformation factors (all these quantities are computed during the forward modelling of the diffraction traveltime functions) the equations above yield the following expressions for the horizontal slowness components at the source and receiver locations:

$$p_S = \pi_S + \frac{\frac{\partial}{\partial \Theta} \Delta t(\Theta, \Phi) + \frac{\partial}{\partial \Phi} \Delta t(\Theta, \Phi)}{2 \frac{\partial x_S}{\partial \gamma_S}} \quad (4.17a)$$

$$p_G = \pi_G + \frac{\frac{\partial}{\partial \Theta} \Delta t(\Theta, \Phi) - \frac{\partial}{\partial \Phi} \Delta t(\Theta, \Phi)}{2 \frac{\partial x_G}{\partial \gamma_G}} \quad (4.17b)$$

To evaluate these expressions along the CRP trajectory all derivatives have to be taken at the geological dip $\Theta = \Theta_0$. In the case of a velocity model which is consistent with the prestack data the horizontal slowness components related to the diffraction traveltime operator and the reflection event in the seismic prestack data coincide along the CRP trajectory as the time dip difference vanishes for all scattering angles. This states that in a velocity model consistent with the data the CRP trajectory is a trajectory of stationary points, i. e., the seismic reflection event and the diffraction traveltime operator are tangent along this trajectory.

4.5 Data example

Finally in this chapter, two data examples are presented to illustrate the above introduced tomographic inversion scheme.

The first data example makes use of only a single pick in a ZO section. This is, in terms of pick distribution, the sparsest example that can be considered. I will not show the inverted velocity model because it can not be assumed that a reasonable model is obtained under such circumstances. However, this example clearly demonstrates the meaning of kinematically fitting a CIG and a CRP gather.

Starting from the same initial model, three different inversions have been performed, each with two iterations. For the first one I used only ZO data and traveltimes residuals picked in a CIG. For the second one I used only ZO data and traveltimes residuals picked in a CRP gather. No time dip residuals have been used in these inversions. In the third one the above presented inversion scheme has been applied. All inversions can be performed with the same implementation: when only traveltimes residuals and ZO data should be used, the time dip residuals and the associated Fréchet derivatives are weighted with zero. The CIGs needed for picking the associated time residuals are directly obtained from the moveout corrected prestack data cubes: all traces with the same scattering angle are summed (summation over the illumination angle) and the summation result is phase-corrected. This phase correction is a half time derivative in the here considered 2D case which corresponds to a multiplication with $i\omega$ in the frequency domain (the complex number is called i and ω denotes the circular frequency).

The inversion results are summarised in Figures 4.8, 4.9, and 4.10, respectively. Each figure shows the CRP gather and the CIG associated with the ZO pick. Additionally, depth migrated common-offset sections are shown. They are centred around the depth position obtained in the corresponding inversion. Note that this position is different in all three inversion results.

The first inversion yields a flat CIG. However, the CRP gather is not flat. Thus, CIG and CRP gather are kinematically not equivalent. This results in different dips of the reflector segment associated with the ZO pick in common offset depth migrated sections.

The second inversion yields a flat CRP gather whereas the CIG shows significant residual moveout. This results not only in different dips for different offsets of the reflector segment associated with the ZO pick, but also the depth location after migration differs from offset to offset.

Applying the inversion scheme presented in this chapter a result is obtained where the CIG and the CRP gather are kinematically equivalent: both are flat. This is due to the use of time dip residuals in the inversion. The minimisation of the time dip residuals leads to a consistent geological dip of the reflector segment. Although the obtained velocity model, depth position, and geological dip are not the true values, a model has been obtained which is kinematically consistent with the prestack data for the considered reflection point. This, of course, only holds for this single considered reflection point.

The second data example uses the synthetic prestack data set already described in Section 3.4. For the here presented example, no random noise was added to the data. A ZO section has been simulated by application of CRS stack method. In this ZO section, 412 pick locations have been automatically selected using the technique presented in Section 3.2.5. The CRS stacked section together with the pick locations is shown in Figure 4.11. For each pick location the quantities characterising the ZO stationary point have been extracted and calculated from the CRS attribute sections. The B-spline representation of the velocity model (see Appendix A) is parameterised by 29 nodes in the horizontal direction with a spacing of 400 m and 18 nodes in the vertical direction with a spacing of 200 m. The velocity at the surface has been set to the true value of 2000 m/s and a constant vertical gradient of 0.67 s^{-1} is initially applied. All pick locations have been mapped to their associated reflection point position in depth by normal ray tracing. In this way, initial estimates of the reflection point position and geological dip related to each pick are obtained.

Six iterations of the above presented inversion scheme have been performed. In all iterations the

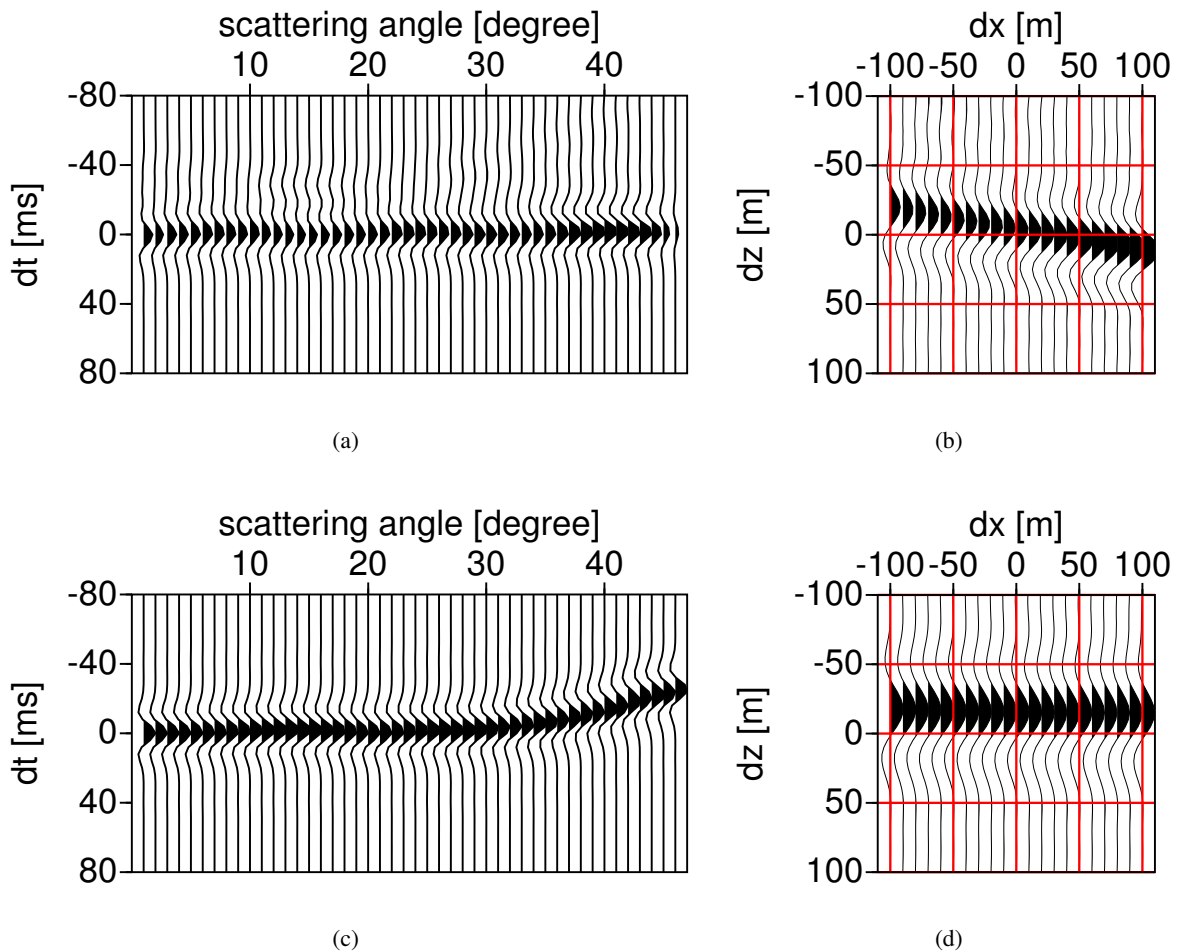


Figure 4.8: Inversion results using traveltime residuals picked in a CIG, no dip residuals have been included in the inversion: (a) flat CIG, (b) depth migrated reflector segment for offset 0 m, (c) CRP gather, (d) depth migrated reflector segment for offset 3800 m. The difference between recording time and diffraction traveltimes is denoted by dt ; dx and dz denote the horizontal and vertical distance to the reflection point position, respectively.

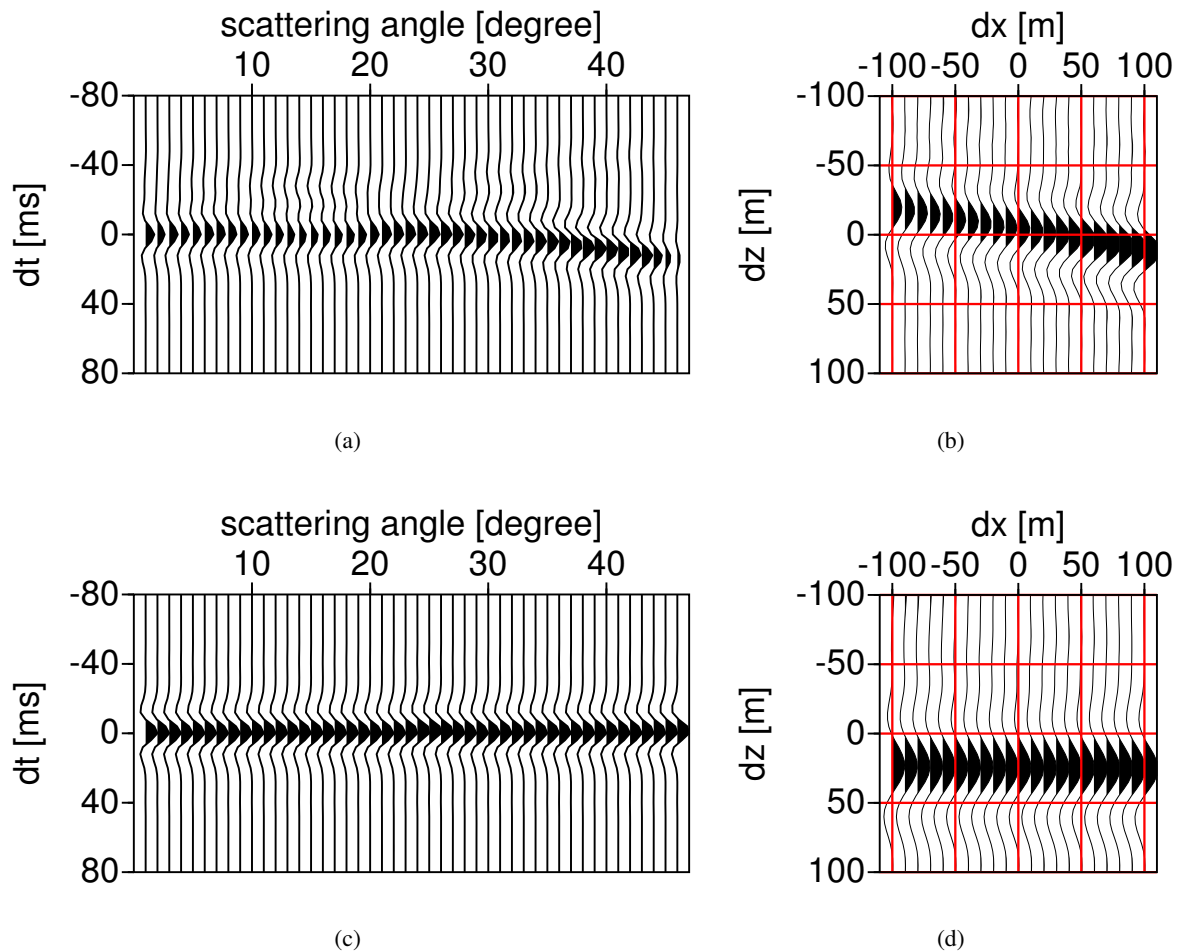


Figure 4.9: Inversion results using traveltime residuals picked in a CRP gather, no dip residuals have been included in the inversion: (a) CIG, (b) depth migrated reflector segment for offset 0 m, (c) flat CRP gather, (d) depth migrated reflector segment for offset 3800 m. The difference between recording time and diffraction traveltimes is denoted by dt ; dx and dz denote the horizontal and vertical distance to the reflection point position, respectively.

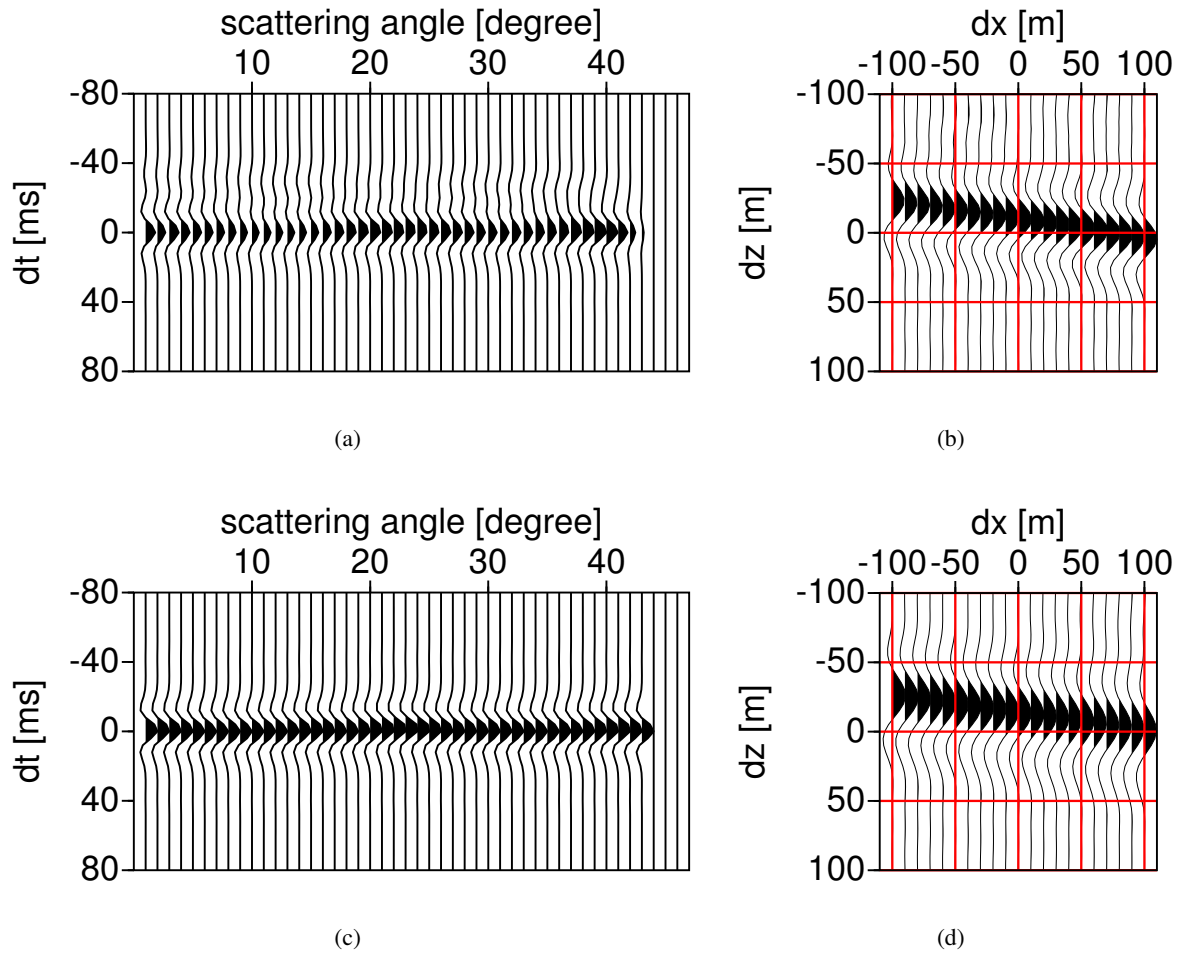


Figure 4.10: Results of the new inversion scheme: (a) flat CIG, (b) depth migrated reflector segment for offset 0 m, (c) flat CRP gather, (d) depth migrated reflector segment for offset 3800 m. The difference between recording time and diffraction traveltimes is denoted by dt ; dx and dz denote the horizontal and vertical distance to the reflection point position, respectively.

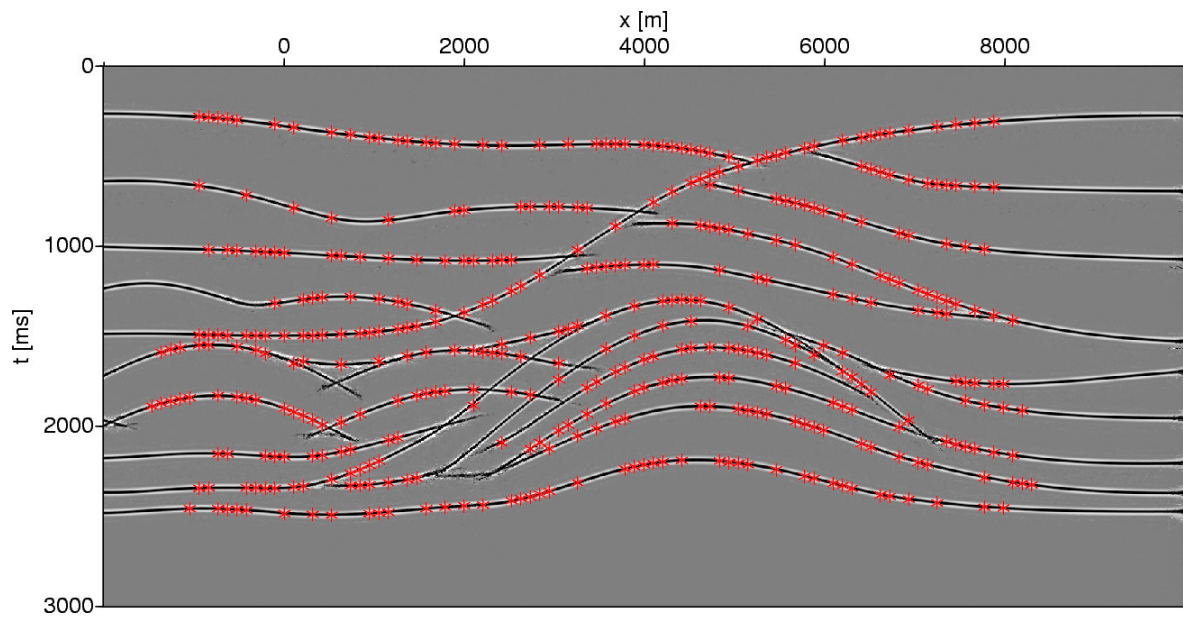


Figure 4.11: ZO stack section with 412 selected pick locations.

depth locations and geological dips obtained as inversion result in the previous iteration have been used, i. e., these quantities have not been recalculated by normal ray tracing after each iteration. Figure 4.12 shows the obtained reflection point positions on top of the true, blocky velocity model used for modelling the prestack data. The obtained depth positions fit well to the reflector structure. The same holds for the stack of prestack depth migrated common-offset sections. Figure 4.14 displays an overlay of the migration result and the true velocity model. Figure 4.13 shows the smooth velocity model obtained by the inversion. On top, the prestack migration result is displayed again. The smooth velocity distribution follows the reflector structure reasonably well although this was not forced by the application of the corresponding additional constraint (see Chapter 5 for details) which minimises the velocity variation along the reflector structure. Generally, a large degree of similarity can be observed comparing the true velocity model and the inversion result. Figure 4.15 shows selected CIGs which are all reasonably flat. For all migration results offsets up to 4000 m have been considered. This maximum offset is twice as large as that used in the data example in Section 3.4 showing the application of NIP-wave tomography. These results demonstrate that the inversion scheme presented in this chapter is suitable to find velocity models which are consistent with the prestack data for a large offset range.

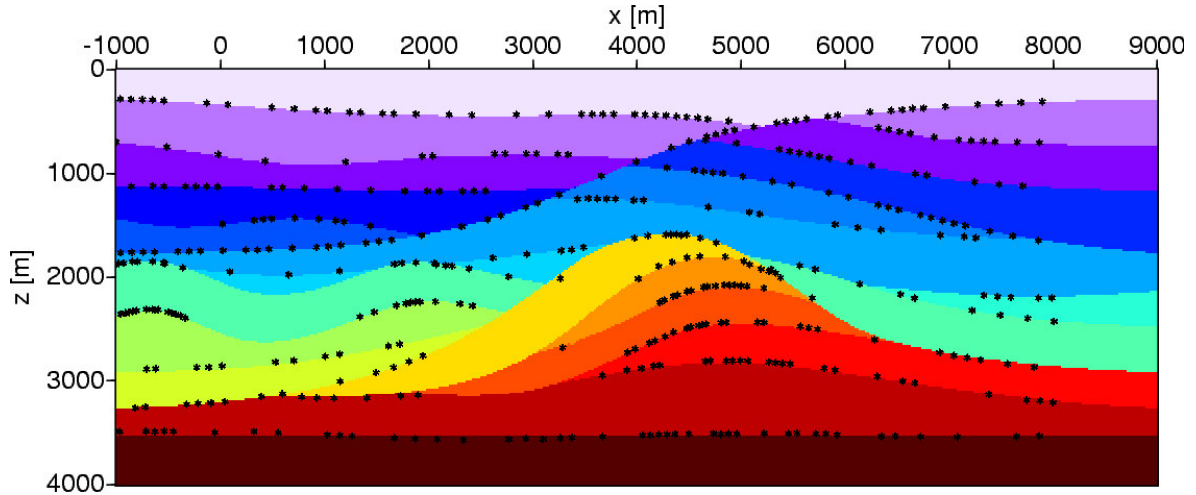


Figure 4.12: Inverted reflection point locations with the true, blocky velocity model in the background.

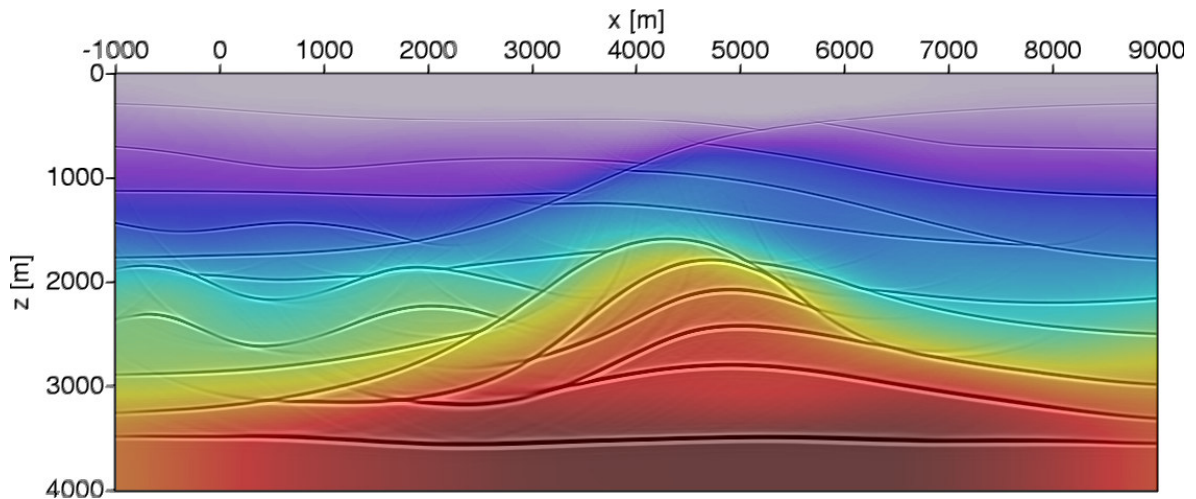


Figure 4.13: Inverted smooth velocity model overlaid with the stack of prestack depth migrated common-offset sections.

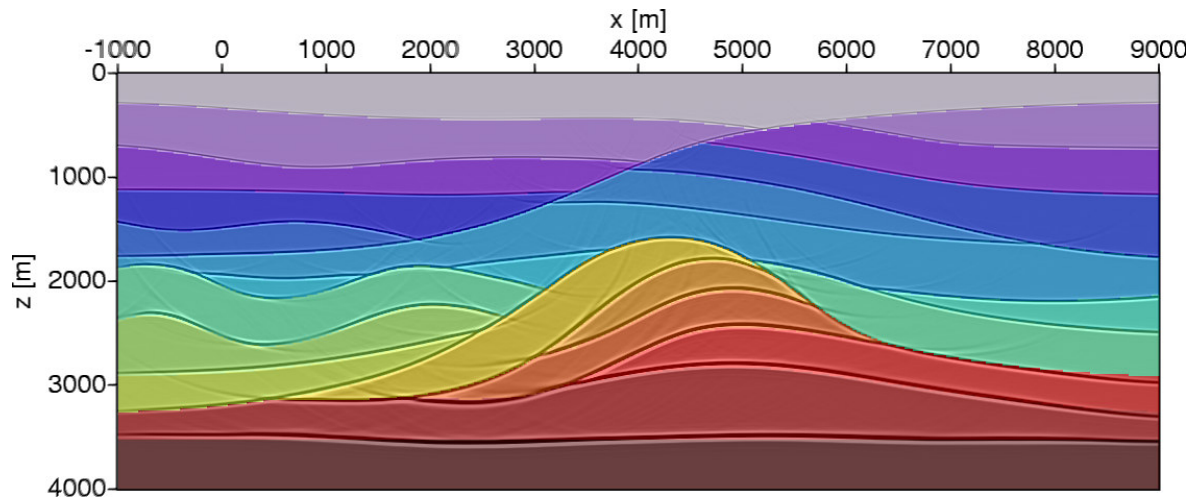


Figure 4.14: The stack of prestack depth migrated common-offset sections using the inverted velocity model with the true, blocky velocity model in the background.

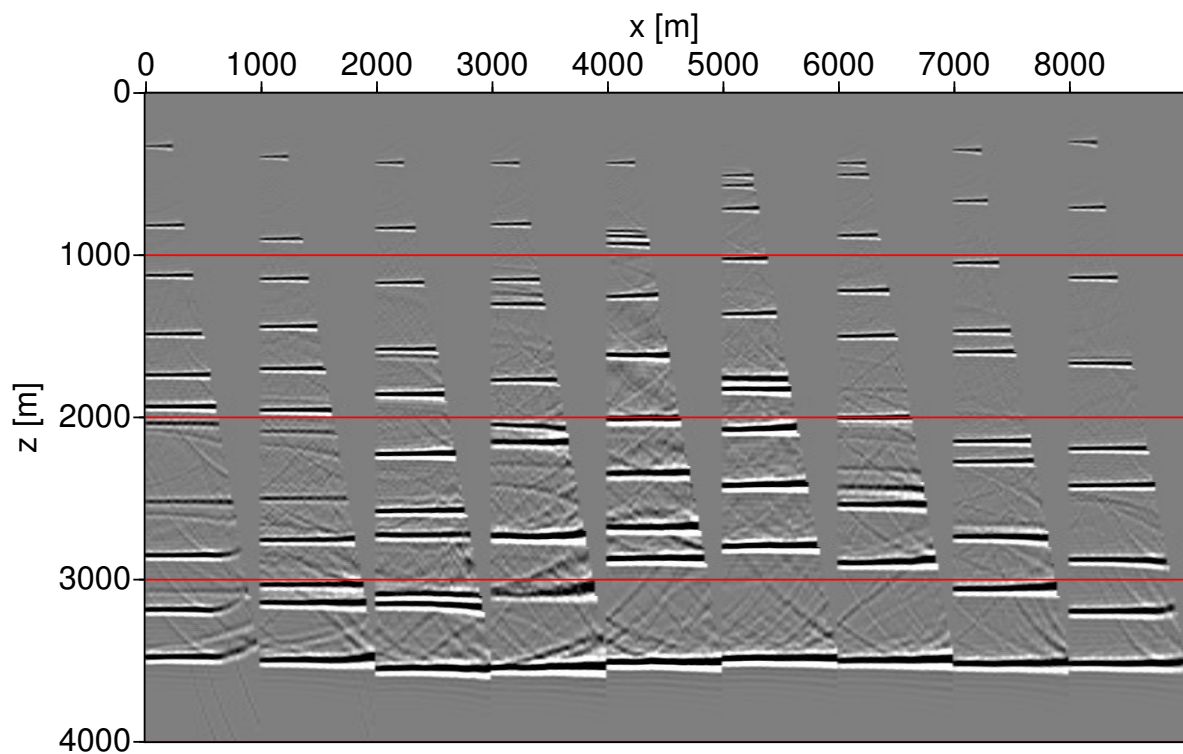


Figure 4.15: Selected CIGs. A maximum offset of 4000 m is displayed.

Chapter 5

Solution of the inverse problem and additional constraints

Both tomographic inversion schemes presented in the previous chapters aim at minimising the misfit between measured and forward-modelled data. This problem is formulated as the minimisation of a cost-function in the least-squares sense. In this chapter, I describe the solution of the inverse problem in a general notation valid for both inversion schemes. This chapter also covers the regularisation necessary for a stable solution as well as additional constraints that may be applied.

Vectors and matrices introduced in this chapter do not follow the general notation outlined in Chapter 1. Their dimensionality depends on the number of data points used in the tomographic inversion schemes.

5.1 The solution of the inverse problem and its regularisation

Let \mathbf{d}_{mea} and \mathbf{d}_{mod} denote the measured and forward-modelled data arranged in a vector. For the individual components building up these vectors the reader is referred to the respective chapters describing the inversion schemes. Using these data vectors the cost-function Ψ is formulated in the following way:

$$\Psi = \|\mathbf{d}_{\text{mea}} - \mathbf{d}_{\text{mod}}\|^2 = \frac{1}{2} \Delta \mathbf{d}^T(\mathbf{m}) \mathbf{C}_D^{-1} \Delta \mathbf{d}(\mathbf{m}), \quad (5.1)$$

where $\Delta \mathbf{d} = \mathbf{d}_{\text{mea}} - \mathbf{d}_{\text{mod}}$ is the data difference vector. The matrix \mathbf{C}_D is sometimes interpreted as a data covariance matrix (Tarantola, 1987). In this thesis it is used in this sense when data derived uncertainties of kinematic wavefield attributes are available. Otherwise, it is assumed to be diagonal and to contain constant weights for each data type. In any case it effectively weights the different data components. Methods for the minimisation of function 5.1 are described in many textbooks on inverse theory (see, e. g., Menke, 1984; Tarantola, 1987). Here, I will follow the lines of Duvaneck (2004a).

Forward-modelled data are obtained in both inversion schemes using dynamic ray tracing. See Chapter 2 for details on the formulation used in this thesis. Since the dynamic ray tracing results depend non-linearly on the subsurface velocity distribution the inverse problem is non-linear, too. The non-linear forward modelling operator yields the modelled data values \mathbf{d}_{mod} :

$$\mathbf{d}_{\text{mod}} = \mathbf{f}(\mathbf{m}), \quad (5.2)$$

where \mathbf{m} is a vector containing all model parameters. Because of the nonlinearity of the inverse problem it can not be solved directly. Instead, the inverse problem is solved iteratively. Suited methods can be found in, e.g., [Menke \(1984\)](#).

The forward modelling operator is linearised during each iteration. Starting with a first guess model $\mathbf{m}^{(0)}$, a sequence of model updates $\Delta\mathbf{m}$ is found such that the process converges to a minimum. Having found a model $\mathbf{m}^{(n)}$ in the n th iteration, the modelling operator \mathbf{f} can locally be linearised expanding it in $\mathbf{m}^{(n)}$ ([Menke, 1984](#)):

$$\mathbf{f}(\mathbf{m}^{(n)} + \Delta\mathbf{m}^{(n)}) \approx \mathbf{f}(\mathbf{m}^{(n)}) + \mathbf{F}^{(n)} \Delta\mathbf{m}^{(n)}. \quad (5.3)$$

The model used for the $(n + 1)$ st iteration is obtained by

$$\mathbf{m}^{(n+1)} = \mathbf{m}^{(n)} + \Delta\mathbf{m}^{(n)}. \quad (5.4)$$

The matrix $\mathbf{F}^{(n)}$ in equation (5.3) contains the Fréchet derivatives, i. e., the first partial derivatives of the components of the current modelled data vector $\mathbf{d}_{\text{mod}}^{(n)}$ with respect to the components of the current model parameter vector $\mathbf{m}^{(n)}$. The Fréchet derivatives are obtained during forward modelling by dynamic ray tracing and ray-perturbation theory. Their calculation is outlined in [Appendix B](#).

A necessary condition for Ψ to have a minimum after the n th iteration is that its gradient with respect to the $(n + 1)$ st model vector $\mathbf{m}^{(n+1)}$ vanishes, i. e.,

$$\nabla_{\mathbf{m}^{(n+1)}} \Psi = \mathbf{0}. \quad (5.5)$$

This implies that the first partial derivatives of the cost function Ψ with respect to all individual components of the model vector $\mathbf{m}^{(n+1)}$ vanish. Under consideration of equation (5.3) the gradient of equation (5.1) reads

$$\nabla_{\mathbf{m}^{(n+1)}} \Psi \approx -\mathbf{F}^{(n)T} \mathbf{C}_D^{-1} (\Delta\mathbf{d}(\mathbf{m}^{(n)}) - \mathbf{F}^{(n)} \Delta\mathbf{m}^{(n)}) = \mathbf{0}. \quad (5.6)$$

Equation (5.6) yields the Gauss-Newton solution for $\Delta\mathbf{m}^{(n)}$ ([Lines and Treitel, 1984](#)) if the inverse of $\mathbf{F}^{(n)T} \mathbf{C}_D^{-1} \mathbf{F}^{(n)}$ exists:

$$\Delta\mathbf{m}^{(n)} = [\mathbf{F}^{(n)T} \mathbf{C}_D^{-1} \mathbf{F}^{(n)}]^{-1} \mathbf{F}^{(n)T} \mathbf{C}_D^{-1} \Delta\mathbf{d}(\mathbf{m}^{(n)}). \quad (5.7)$$

In practice, equation (5.7) is not directly applicable, because the matrix $\mathbf{F}^{(n)T} \mathbf{C}_D^{-1} \mathbf{F}^{(n)}$ is usually ill-conditioned. This means that not all model components are sufficiently constrained by the data alone. There might be subsurface regions where no reflection points are located. This difficulty is removed by introducing additional a priori information to further constrain the model and to regularise the problem ([Delprat-Jannaud and Lailly, 1993](#)). This is done by minimising the second derivatives of the velocity model which means to minimise its curvature. Thus, the simplest model is searched for which is consistent with the data. In order to be as independent as possible of the discrete structure of the grid of B-spline nodes, the minimum curvature condition is applied on the smooth velocity model itself and not on the B-spline coefficients. The minimum curvature criterion is formulated here for the general 3D case. The corresponding equations for the 2D case are obtained by omitting all derivatives with respect to the y -coordinate.

The second spatial derivative of the velocity, described by B-splines, with respect to the x -coordinate reads

$$\frac{\partial^2 v(x, y, z)}{\partial x^2} = \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} \sum_{k=1}^{n_z} v_{ijk} \frac{\partial^2 \beta_i(x)}{\partial x^2} \beta_j(y) \beta_k(z). \quad (5.8)$$

See Appendix A for details on the velocity model parameterisation. The partial derivatives with respect to y and z are given by analogous expressions. The L_2 -norm of (5.8) reads

$$\left\| \frac{\partial^2 v}{\partial x^2} \right\|_2^2 = \int_x \int_y \int_z \left(\frac{\partial^2 v(x, y, z)}{\partial x^2} \right)^2 dz dy dx = \mathbf{m}_{(v)}^T \mathbf{D}^{(xx)} \mathbf{m}_{(v)}, \quad (5.9)$$

where $\mathbf{m}_{(v)}$ denotes that part of the model parameter vector containing the B-spline coefficients. Similar expressions are derived for the second derivatives with respect to y and z . Mixed derivatives are not considered in the regularisation. This leads to the following term to be included in the cost function Ψ :

$$\int_x \int_y \int_z \epsilon^{xx} \left(\frac{\partial^2 v}{\partial x^2} \right)^2 + \epsilon^{yy} \left(\frac{\partial^2 v}{\partial y^2} \right)^2 + \epsilon^{zz} \left(\frac{\partial^2 v}{\partial z^2} \right)^2 + \epsilon v^2(x, z) dz dy dx = \mathbf{m}_{(v)}^T \mathbf{D}'' \mathbf{m}_{(v)}. \quad (5.10)$$

The weighting factors ϵ^{xx} , ϵ^{yy} , ϵ^{zz} , and ϵ are used for normalisation and to balance the contributions of the different terms. Each of them can be chosen spatially variable (Duvneek, 2004a). This spatially dependent regularisation has not been used in the examples presented in this thesis as it is difficult to control in practice. It is not further described. However, it might be useful at the edges of the velocity model and in case of a strongly non-uniform ray coverage of the subsurface.

Including the term (5.10), the cost function to be minimised is given by (compare to equation (5.1))

$$2\Psi(\mathbf{m}) = \Delta \mathbf{d}^T(\mathbf{m}) \mathbf{C}_D^{-1} \Delta \mathbf{d}(\mathbf{m}) + \hat{\epsilon} \mathbf{m}_{(v)}^T \mathbf{D}'' \mathbf{m}_{(v)}, \quad (5.11)$$

where $\hat{\epsilon}$ is the overall weighting of the regularisation term. For the further calculations a matrix \mathbf{B} with $\mathbf{B}^T \mathbf{B} = \hat{\epsilon} \mathbf{D}''$ has to be found (Duvneek, 2004a). This is only possible if \mathbf{D}'' is positive definite, i. e., it fulfils the following condition for all possible velocity models:

$$\mathbf{m}_{(v)}^T \mathbf{D}'' \mathbf{m}_{(v)} > 0, \quad \text{for all } \mathbf{m}_{(v)} \neq \mathbf{0}. \quad (5.12)$$

To account for this condition, the term $\epsilon v^2(x, y, z)$ is included in equation (5.10). Considering only second derivatives is generally not sufficient. For example, in the case of a constant velocity model equation (5.12) trivially fails. Because there is no physical reason for minimising the velocity itself, ϵ is chosen much smaller than the other weighting factors.

Assuming that the forward modelling operator can locally be linearised around $\mathbf{m}^{(n)}$ as in equation (5.3), one obtains for the gradient of the cost-function

$$\begin{aligned} \nabla_{\mathbf{m}^{(n+1)}} \Psi &= -\mathbf{F}^{(n)T} \mathbf{C}_D^{-1} (\mathbf{d}_{\text{mea}} - \mathbf{f}(\mathbf{m}^{(n+1)})) + \hat{\epsilon} \mathbf{D}'' \mathbf{m}_{(v)}^{(n+1)} \\ &\approx -\mathbf{F}^{(n)T} \mathbf{C}_D^{-1} \Delta \mathbf{d}(\mathbf{m}^{(n)}) + \mathbf{F}^{(n)T} \mathbf{C}_D^{-1} \mathbf{F}^{(n)} \Delta \mathbf{m}^{(n)} + \\ &\quad \hat{\epsilon} \mathbf{D}'' (\mathbf{m}_{(v)}^{(n)} + \Delta \mathbf{m}^{(n)}). \end{aligned} \quad (5.13)$$

The necessary condition $\nabla_{\mathbf{m}^{(n+1)}} \Psi = \mathbf{0}$ for a minimum of the cost function (5.11) after the n th iteration leads to (Duvneek, 2004a)

$$\hat{\mathbf{F}}^{(n)T} \hat{\mathbf{F}}^{(n)} \Delta \mathbf{m}^{(n)} = \hat{\mathbf{F}}^{(n)} \Delta \hat{\mathbf{d}}^{(n)}, \quad (5.14)$$

with

$$\hat{\mathbf{F}}^{(n)} = \begin{pmatrix} \mathbf{C}_D^{-\frac{1}{2}} \mathbf{F}^{(n)} \\ [\mathbf{0}, \mathbf{B}^{(n)}] \end{pmatrix}, \quad \Delta \hat{\mathbf{d}}^{(n)} = \begin{pmatrix} \mathbf{C}_D^{-\frac{1}{2}} \Delta \mathbf{d}(\mathbf{m}^{(n)}) \\ -[\mathbf{0}, \mathbf{B}^{(n)}] \mathbf{m}^{(n)} \end{pmatrix}. \quad (5.15)$$

Therefore, the searched-for model update $\Delta \mathbf{m}^{(n+1)}$ is the least-squares solution of

$$\hat{\mathbf{F}}^{(n)} \Delta \mathbf{m}^{(n)} = \Delta \hat{\mathbf{d}}^{(n)} . \quad (5.16)$$

The updated model vector is computed according to

$$\mathbf{m}^{(n+1)} = \mathbf{m}^{(n)} + \lambda \Delta \mathbf{m}^{(n)} , \quad (5.17)$$

with $0 < \lambda \leq 1$. This means that not the complete model update is applied. This accounts for the applied linearisations. Due to the non-linearity of the inverse problem, the cost-function is generally not a quadratic function of \mathbf{m} . In that case a non-iterative inversion scheme could be applied (Tarantola, 1987).

5.2 Additional constraints

In addition to the input data and the minimisation of curvature, further criteria may be used to constrain the velocity model. Often, a priori information about the velocity model from geology or well-log data is available. In many geological settings it is plausible to assume only small variations of the velocity along reflectors. In the following, strategies are described to include such a priori information in the inversion process.

5.2.1 A-priori velocity information

The more reliable a-priori information about the subsurface velocity model is available and included into the inversion scheme the more stable becomes the tomographic inversion process. A priori information about the velocity model may stem from several sources: direct measurements, boreholes or geological knowledge. The probably most obvious case is marine seismic data where the near-surface velocity has to coincide with the acoustic wave-propagation velocity in water.

The a priori velocity values are part of the data components in both inversion schemes presented in the previous chapters. Assuming that a priori velocity values are available at n_{vdata} locations, a total of n_{vdata} elements

$$v(x_i, y_i, z_i) , \quad i = 1, \dots, n_{\text{vdata}} , \quad (5.18)$$

is appended to the data vector \mathbf{d}_{mea} . During the inversion the misfit between these a priori velocity values and their forward-modelled counterparts is minimised. Each of these additional data points has its own weight in the matrix \mathbf{C}_D . Note that the a priori known velocity values do not depend on that part of the model description related to the NIP-locations. Therefore, that part of matrix $\hat{\mathbf{F}}$ is padded with zeros.

5.2.2 Minimum velocity variation along reflectors

In some geological settings it is reasonable to assume that the velocity variation is much stronger in the direction normal to the reflector than along the reflecting interface. However, as both inversion schemes presented in previous chapters aim at finding a smooth background velocity model suitable for depth migration this criterion should be used with care.

As the model description does not include any continuous reflectors the criterion of minimum velocity variation along reflection interfaces can only be locally applied at the NIP of the normal ray associated with each pick. At these positions a unit tangent vector $\underline{\hat{\mathbf{e}}}$ to the normal ray can be constructed using the current model parameters. For the 3D case it is given by

$$\underline{\hat{\mathbf{e}}} = \left(e_x, e_y, \sqrt{1 - e_x^2 - e_y^2} \right)^T . \quad (5.19)$$

The normalised velocity gradient at the NIP must be collinear to this unit tangent vector. As the ray is normal to the interface at the NIP this implies a minimum local velocity variation along the reflector. The normalised velocity gradient has to be used because otherwise the criterion would depend on the magnitude of local velocity changes. The normalised velocity gradient for the general 3D case is expressed by

$$\underline{\mathbf{e}}^{(v)} = \frac{\nabla v}{|\nabla v|} = \frac{1}{|\nabla v|} \left(\frac{\partial v}{\partial x}, \frac{\partial v}{\partial y}, \frac{\partial v}{\partial z} \right)^T . \quad (5.20)$$

The minimisation of the local velocity variation along the reflector is the analogue to the minimisation of

$$1 - \underline{\hat{\mathbf{e}}} \cdot \underline{\mathbf{e}}^{(v)} \quad (5.21)$$

for each NIP in the model description.

The practical implementation follows a different strategy. Let $\underline{\hat{\mathbf{e}}}_1$ and $\underline{\hat{\mathbf{e}}}_2$ denote two unit vectors in the plane tangent to the reflector at the NIP which form a mutually orthogonal triplet of unit vectors with $\underline{\hat{\mathbf{e}}}$. The projection of the velocity gradient onto the plane locally tangent to the reflector, i. e., normal to the ZO ray, is given for the 3D case by

$$\nabla_{\perp} v = \left(\underline{\hat{\mathbf{e}}}_1 \cdot \underline{\mathbf{e}}^{(v)}, \underline{\hat{\mathbf{e}}}_2 \cdot \underline{\mathbf{e}}^{(v)} \right)^T . \quad (5.22)$$

The local velocity structure at the NIP locations is constraint in the inversion by minimising

$$|\nabla_{\perp} v|_i = \left[\left(\underline{\hat{\mathbf{e}}}_1 \cdot \underline{\mathbf{e}}^{(v)} \right)^2 + \left(\underline{\hat{\mathbf{e}}}_2 \cdot \underline{\mathbf{e}}^{(v)} \right)^2 \right]_i^{\frac{1}{2}}, \quad i = 1, \dots, n_{\text{data}} . \quad (5.23)$$

This makes another n_{data} elements in the data vector \mathbf{d}_{mea} . Of course, this criterion can only be applied if the pick locations which are input to the tomography actually correspond to a reflection point in depth. An application for diffraction events makes no sense. Due to this fact and the smooth character of the velocity models obtained by the inversion schemes, the weights in matrix \mathbf{C}_D should be chosen with care: a large error should be allowed relative to that for the data points and other velocity constraints.

Chapter 6

Applications

In this chapter, applications of the inversion methods presented in the previous chapters are presented. I start with a synthetic complex 2D data example to demonstrate the complete work flow from the CRS stack to the update using model-based diffraction traveltime functions. Then, NIP-wave tomography is demonstrated for a 3D real data example with special emphasis placed on a comparison between inversion results obtained with complete azimuth-dependent second traveltime derivatives and second traveltime derivatives in a single azimuthal direction.

6.1 2D synthetic data example

In this section, I present the results obtained with the successive application of both inversion methods presented in this thesis to a complex 2D synthetic data example.

6.1.1 Data description

The synthetic data set was forward-modelled using a Kirchhoff integral-based modelling technique. The data set only contains primary P-P reflection events. It consists of 59361 traces distributed on 981 CMP gathers. Each trace has 1251 samples with a sampling interval of 4 ms. The midpoint spacing is 12.5 m, the maximum offset 3500 m. The smooth velocity distribution used for modelling of the prestack data is displayed along with a structural image in Figure 6.1. The complex overburden makes it difficult to properly image the target area which is centred around $x = 6.5$ km and $z = 2.8$ km. The complexity of the velocity model leads to strongly non-hyperbolic moveout of the reflection events in several CMP gathers. The moveout is even negative at some locations. Figure 6.2 demonstrates the complex moveout behaviour for a part of the CMP gather at $x = 6.95$ km.

6.1.2 CRS stack results

The velocity model estimation work flow started with the application of the CRS stack method. The implementation described in Mann (2002) has been used. Of special importance was an proper choice of the attribute search aperture in the offset dimension in order to reflect the complex velocity model

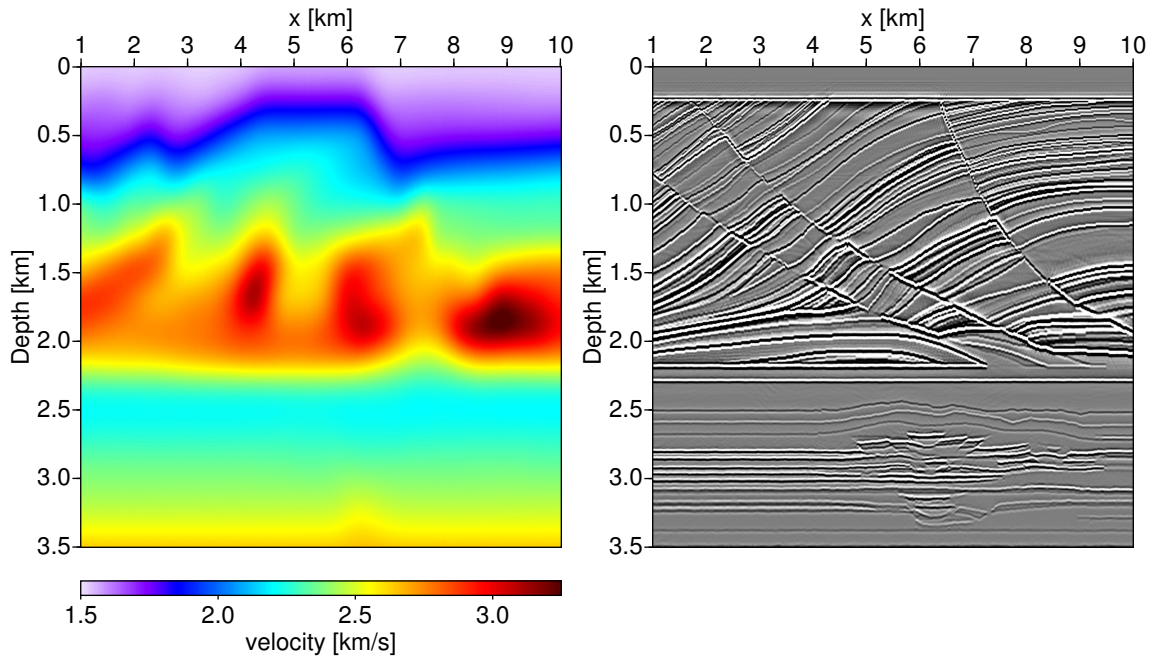


Figure 6.1: Left: smooth velocity model used for forward-modelling of the prestack data. Right: structural image obtained by depth migration of the forward-modelled ZO section.

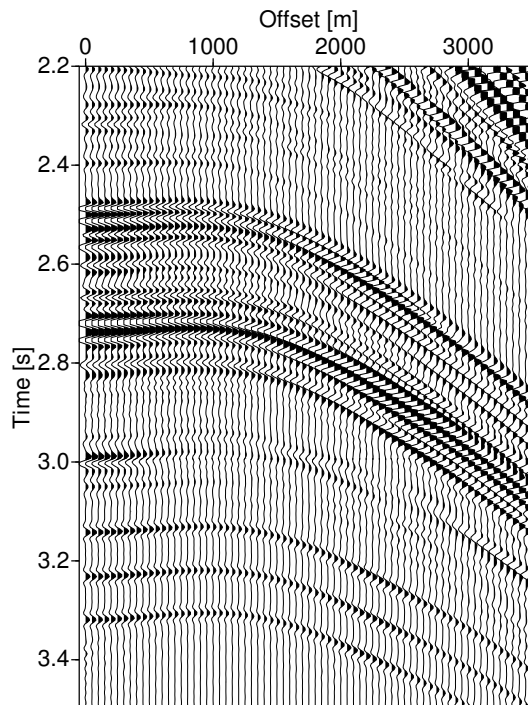


Figure 6.2: CMP gather at $x = 6.95$ km. The complex moveout behaviour at large traveltimes due to the complex velocity model can be clearly observed.

in the CRS stacking parameters. The chosen maximum offset was linearly increasing from 100 m at $t = 0$ s to 2000 m at $t = 2$ s. For larger two-way traveltimes the maximum offset was kept constant. This aperture choice ensured stable attribute search results and allowed to detect the negative moveout present in several CMP gathers. Due to the high-quality data no three-parameter optimisation and smoothing of the CRS stack parameters was performed.

The CRS stack results are displayed in Figure 6.3. Except of the region around $x = 6.5$ km high semblance values are obtained along the CRS stacking operator for most reflection events. The negative moveout present in CMP gathers in that region has been detected by the CRS stack method. The negative moveout corresponds to negative values of the NIP-wavefront curvature radius. They are observed in the corresponding section in Figure 6.3 as coherent white regions for traveltimes larger than 2 s around $x = 6.5$ km. Note that the legend starts at $R_{\text{NIP}} = 0$ km. Thus, the negative values out of the legend's range and coloured white. The high-velocity lenses in the velocity model yield strongly varying NIP wavefront curvature radius values for the reflection events at large traveltimes. The normal ray emergence angles are mostly restricted to $-15^\circ < \beta < 15^\circ$. Only the reflection events associated with the fault planes in the upper part of the model and diffraction events show stronger dips.

6.1.3 NIP-wave tomography and update

The CRS results displayed in Figure 6.3 were input to the automated picking strategy. Valid pick locations had at least a semblance value of 0.45. Picking was performed on every 20th trace in the ZO stack section. The selected 442 pick locations are displayed on top of the CRS stack section in Figure 6.4. No further pick editing was applied.

The CRS attributes of the selected 442 pick locations provided the input for NIP-wave tomography. The velocity model was parameterised with 41 nodes in the x -direction and 27 nodes in the z -direction. The initial B-spline coefficients followed a vertical gradient model with $v = 1500$ m/s at the surface and a gradient of 0.3 s^{-1} . Each input data type was weighted with a constant weight: $w_T = 1$, $w_M = 1$, $w_p = 0.5$, and $w_x = 1$. The surface velocity was constrained to 1500 m/s. The corresponding weight was set to $w_v = 2$. The initial weights for the regularisation were specified with $\epsilon^{xx} = 0.5$, $\epsilon^{zz} = 0.1$, and $\epsilon = 10^{-12}$. These weights were decreased from iteration to iteration according to the square root of the ratio between the new and the old cost-function values. The inversion terminated after ten iterations as the cost-function was no longer decreasing.

The final model obtained with NIP-wave tomography as well as the final estimates of the reflection point locations and geological dips were input to the tomographic inversion scheme using model-based diffraction traveltime functions. This inversion method was applied as an updating technique to further improve the results of NIP-wave tomography. Seven iterations of the inversion scheme described in Chapter 4 were performed. The regularisation weights were set to $\epsilon^{xx} = 0.05$ and $\epsilon^{zz} = 0.01$ for the iterations one and two. These values approximately are twice as large as the final regularisation weights applied in NIP-wave tomography. They were decreased to $\epsilon^{xx} = 0.01$ and $\epsilon^{zz} = 0.005$ for the iterations three, four, and five. Starting with iteration six regularisation weights $\epsilon^{xx} = 0.005$ and $\epsilon^{zz} = 0.0025$ were used. In all iterations $\epsilon = 10^{-12}$ was applied. The ZO data components were weighted with the same constant factors as applied in NIP-wave tomography. The weights for the finite scattering angle time and time dip residuals were set to $w_T^{FO} = 1$ and $w_p^{FO} = 0.1$. The lower weight for the time dip residuals was necessary to keep all data misfits and Fréchet derivatives at a numerically comparable size.

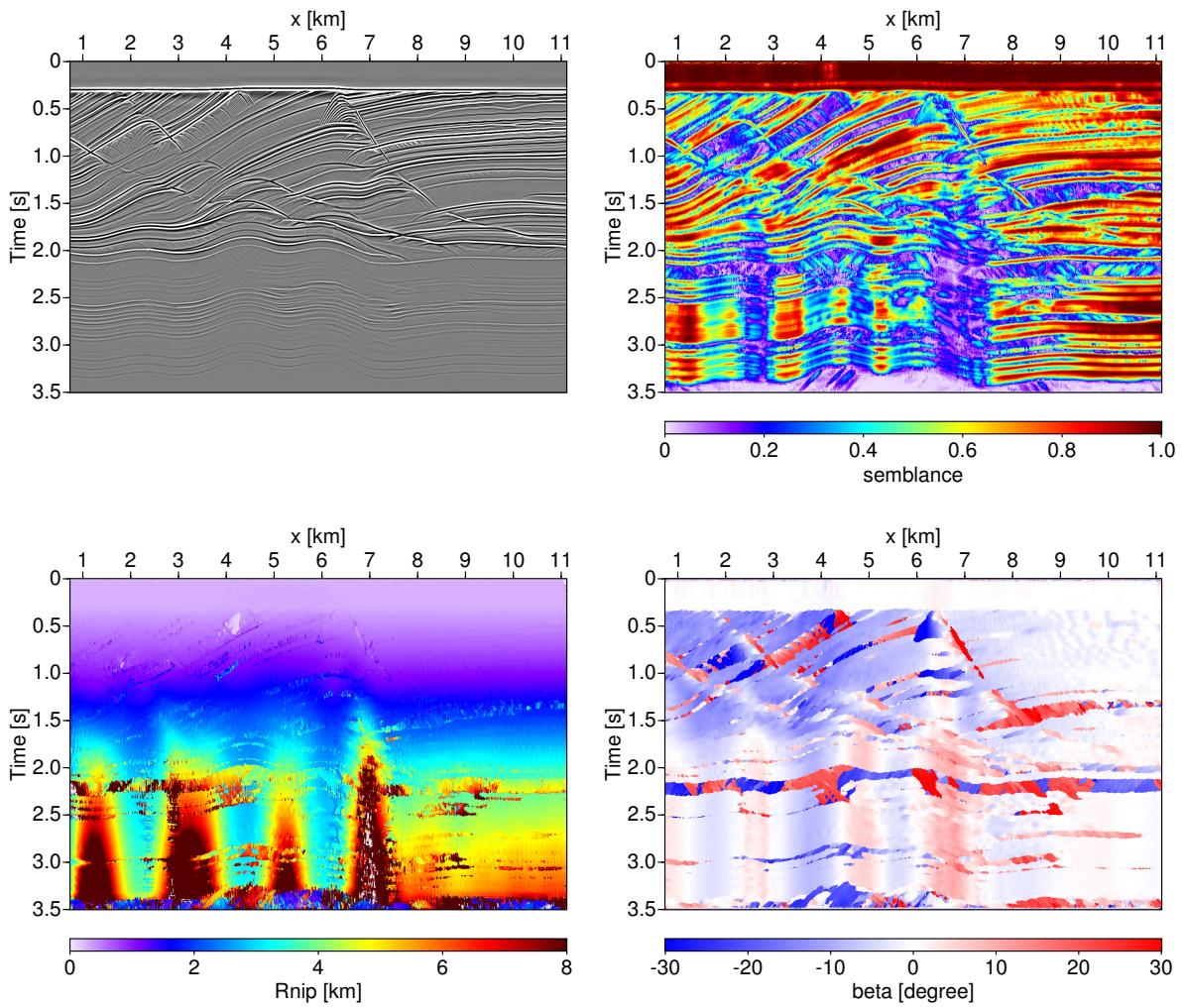


Figure 6.3: CRS stack results. Upper row: CRS stack section (left) and coherence section (right). Lower row: NIP wavefront radius of curvature (left) and normal ray emergence angle (right).

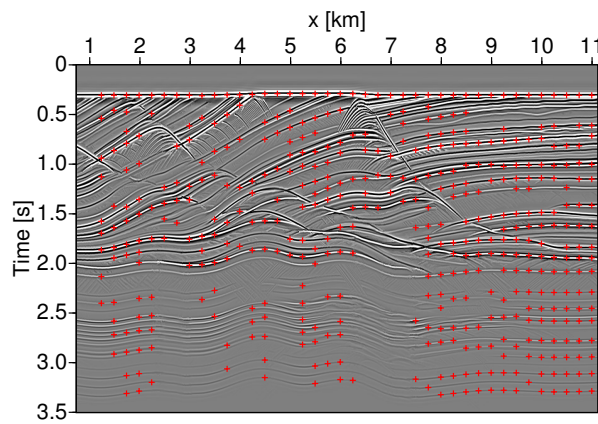


Figure 6.4: Automatically detected pick locations on top of the CRS stack section.

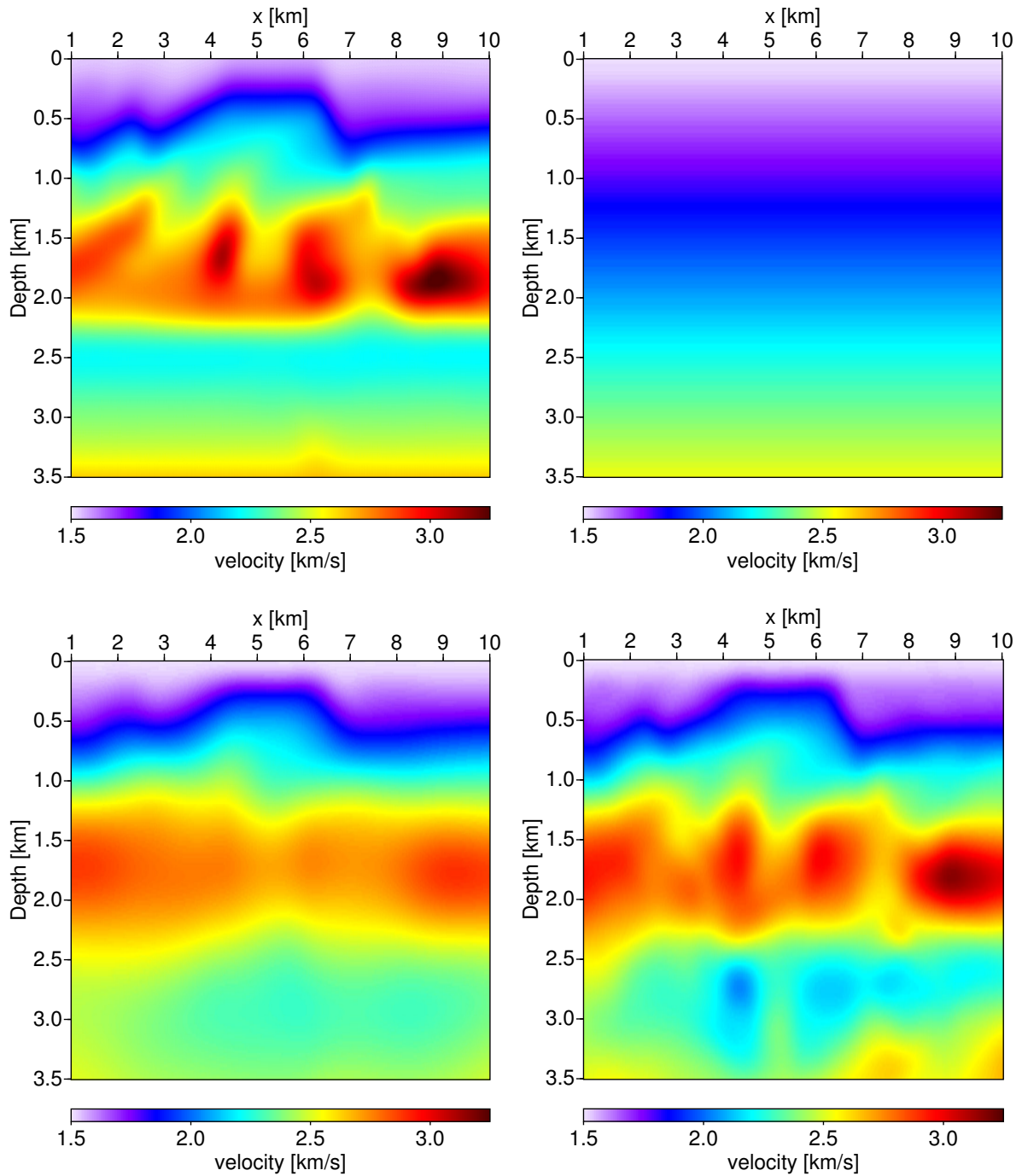


Figure 6.5: Upper row: the true velocity model (left) and the initial velocity model for NIP-wave tomography (right). Lower row: Final velocity models of NIP-wave tomography (left) and the update (right) using model-based diffraction traveltimes.

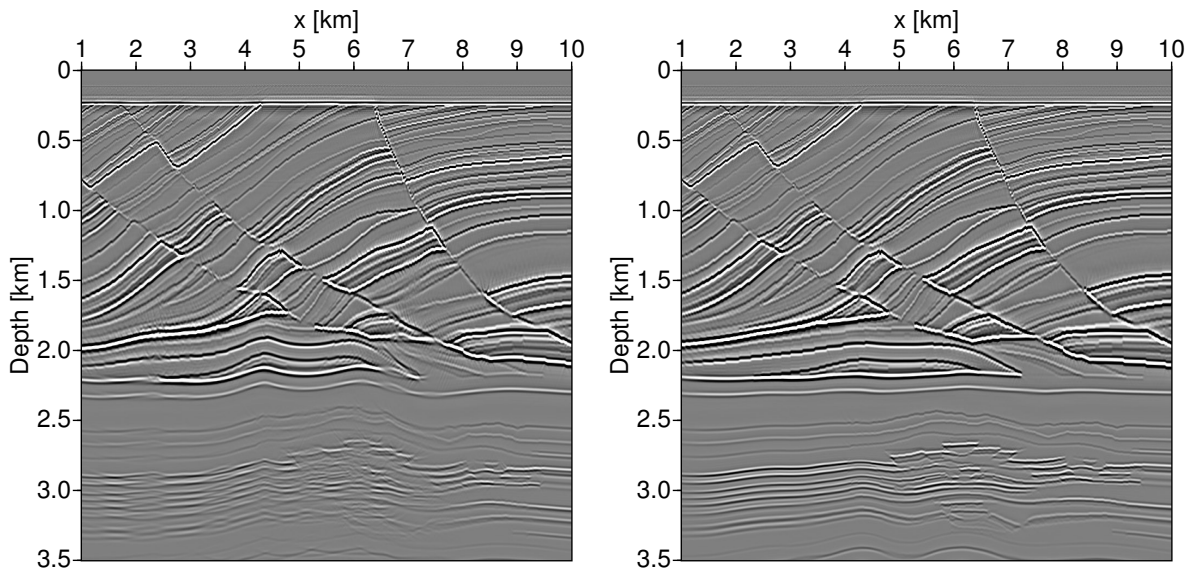


Figure 6.6: Stack of prestack depth migrated common-offset sections. Migration has been performed using the final model of NIP-wave tomography (left) and the model after the update (right) with model-based diffraction traveltimes functions.

In Figure 6.5 the inversion results are depicted along with the smooth velocity model used in forward-modelling of the prestack data and the initial velocity model for NIP-wave tomography. The final model obtained with NIP-wave tomography shows increasing velocities until $z \approx 2$ km. Then the velocity decreases and increases again. This overall structure is also found in the true velocity model. The NIP-wave tomography was not able to resolve the high-velocity lenses present in the true velocity model. However, NIP-wave tomography provided an excellent initial velocity model for the inversion method using model-based diffraction traveltimes operators. The update of the NIP-wave tomography result yields much more details in the velocity model structure. The maximum values of the high-velocity lenses have not been reached. It seems that the absence of high-velocity peaks is compensated for with the inclusion of low velocity lenses. These low velocity zones laterally appear at the same positions than the high velocity zones and are vertically just below. The overall velocity model much closer resembles the true one than the NIP-wave tomography result. This holds especially for the upper part.

6.1.4 Depth migration results

To validate the tomography results Kirchhoff prestack depth migration was performed using the obtained velocity models. Figure 6.6 shows the stack of depth migrated common-offset sections. Maximum offsets equal to depth have been used in the stack. Obviously, the more detailed velocity model after the update leads to a structural image of the subsurface which is much closer to the true one displayed in Figure 6.1. The reflector at $z \approx 2.3$ km is nearly horizontal and the target area below shows highly increased focusing of scattered energy. The deepest reflectors show strong structural distortions even after the update. However, this has to be expected in global tomographic inversion schemes as this is the region constrained by the smallest amount of input data. The selected CIGs displayed in Figure 6.7 clearly demonstrate the increased consistency of the updated velocity model with the prestack data: nearly all CIGs show increased flatness compared to those obtained with the

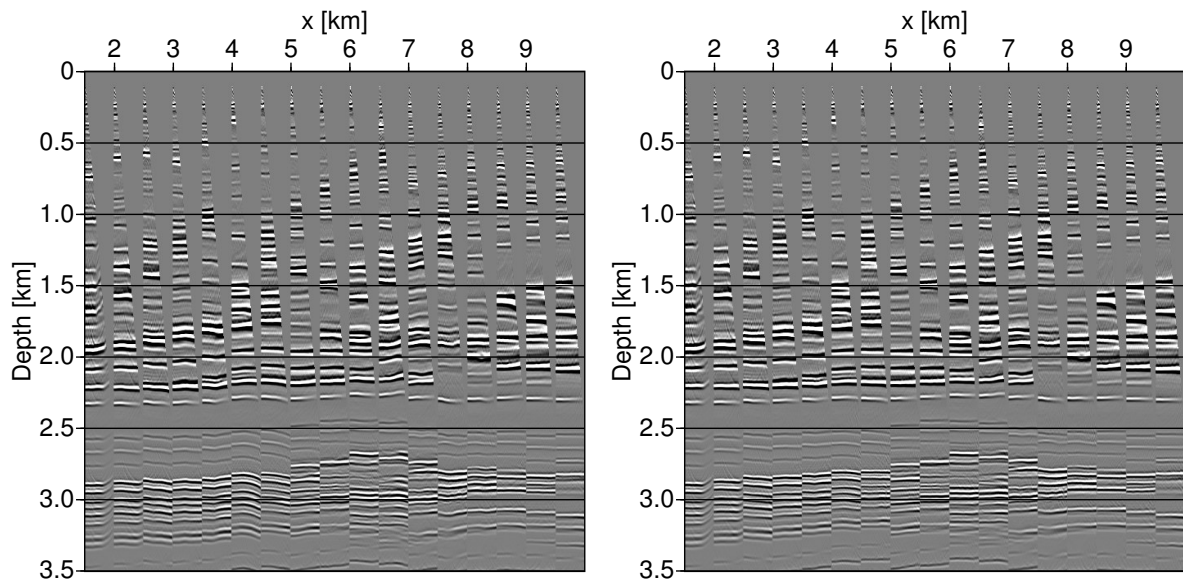


Figure 6.7: CIG with a spacing of 500 m obtained with the final velocity model of NIP-wave tomography (left) and the updated velocity model (right). The maximum offset is 3500 m.

final velocity model of NIP-wave tomography.

6.1.5 Conclusions

I have demonstrated the application of NIP-wave tomography followed by the update technique to a complex 2D synthetic data example. The moveout in the CMP gathers of the prestack data is too complex to be accurately described by the hyperbolic traveltime operator used in the CRS stack method. Therefore, NIP-wave tomography leads to velocity models which are not consistent with the prestack data. However, the input to the tomographic inversion has been minimised. The application of the inversion method using model-based diffraction traveltime functions significantly improves the final velocity model and leads to increased consistency with the prestack data. The final structural image of the subsurface has been significantly improved.

6.2 3D real data example

The inversion scheme based on second-order diffraction traveltime functions is applied to a 3D real land data set in this section. The data set was made available within the project described in [Trappe et al. \(2005\)](#).

The data were acquired with full azimuthal coverage. This allows the usage of azimuth-dependent second-order traveltime derivatives in the inversion. In this section I present a comparison of results obtained with an inversion using second-order traveltime derivatives in just one azimuth and using the complete matrix of second traveltime derivatives. The two inversions are named single-azimuth and multi-azimuth inversion, respectively.

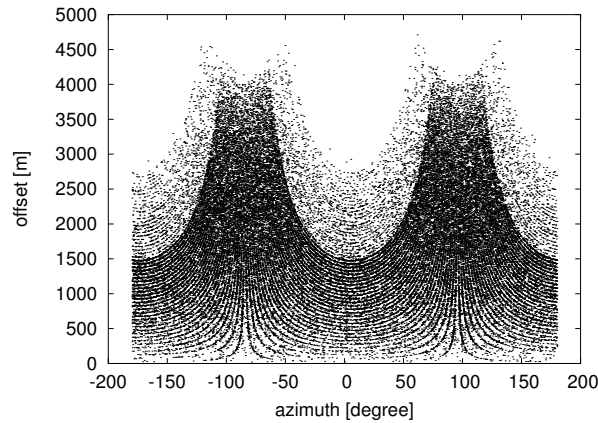


Figure 6.8: Azimuthal distribution of offsets. Displayed is every 25th trace of the prestack data set. Angles are measured counterclockwise starting from the positive x-axis.

6.2.1 Data description

The offset vectors in each CMP volume cover the complete azimuthal range. However, the maximum offset differs from azimuth to azimuth and varies between 2700 m and 4700 m with smaller values at the edges of the prestack data volume. The largest offsets are found with directions nearly parallel to the y-axis defined in Figure 6.9. The azimuthal distribution of offset vectors is shown in Figure 6.8 for every 25th trace of the prestack data set. The complete data set has 1610552 traces with 1001 samples and a sampling rate of 4 ms. The traces are distributed on 51192 CMP bins with a mean fold of approximately 32 traces. The CMP locations have a mean spacing of 25 m. Their spatial distribution is displayed in Figure 6.9.

6.2.2 CRS stack

The first step in the data processing chain was the application of the CRS stack using the implementation described in Müller (2003). The final step in the CRS search strategy was a simultaneous optimisation of all eight stacking parameters. Figure 6.9 shows the location of all traces in the stacked volume. Highlighted in red are two cross sections for which the CRS results are depicted in the following. All Figures show sections of the volumes used as input for the tomographic inversion. The attribute volumes have been smoothed in an event-consistent way according to the algorithm outlined in Section 3.1.3: a combination of mean and median filtering has been used to remove outliers and unphysical fluctuations. The final CRS stacked volume has been computed using these smoothed stacking parameters.

The results presented in this section, including the tomographic inversion and depth migration, are restricted to the right part of the stacked ZO volume. In the left part, the surface topography is undulating much more than in the right part. These undulations have strong effects on the CRS attributes and, thus, the attributes would distort the tomographic inversion. Although undulations of the surface topography are present in the right part as well, they have been neglected throughout the processing. This should be kept in mind when the results are interpreted.

Figure 6.10 shows the final CRS stack. The semblance value obtained along the CRS stacking

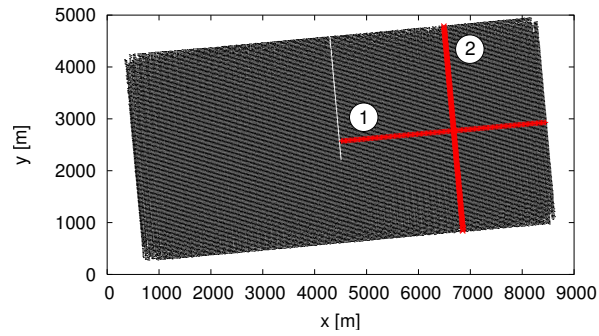


Figure 6.9: ZO trace locations. Highlighted in red are two cross sections for which the CRS results are shown. They are referred to as profile 1 and profile 2 in the text.

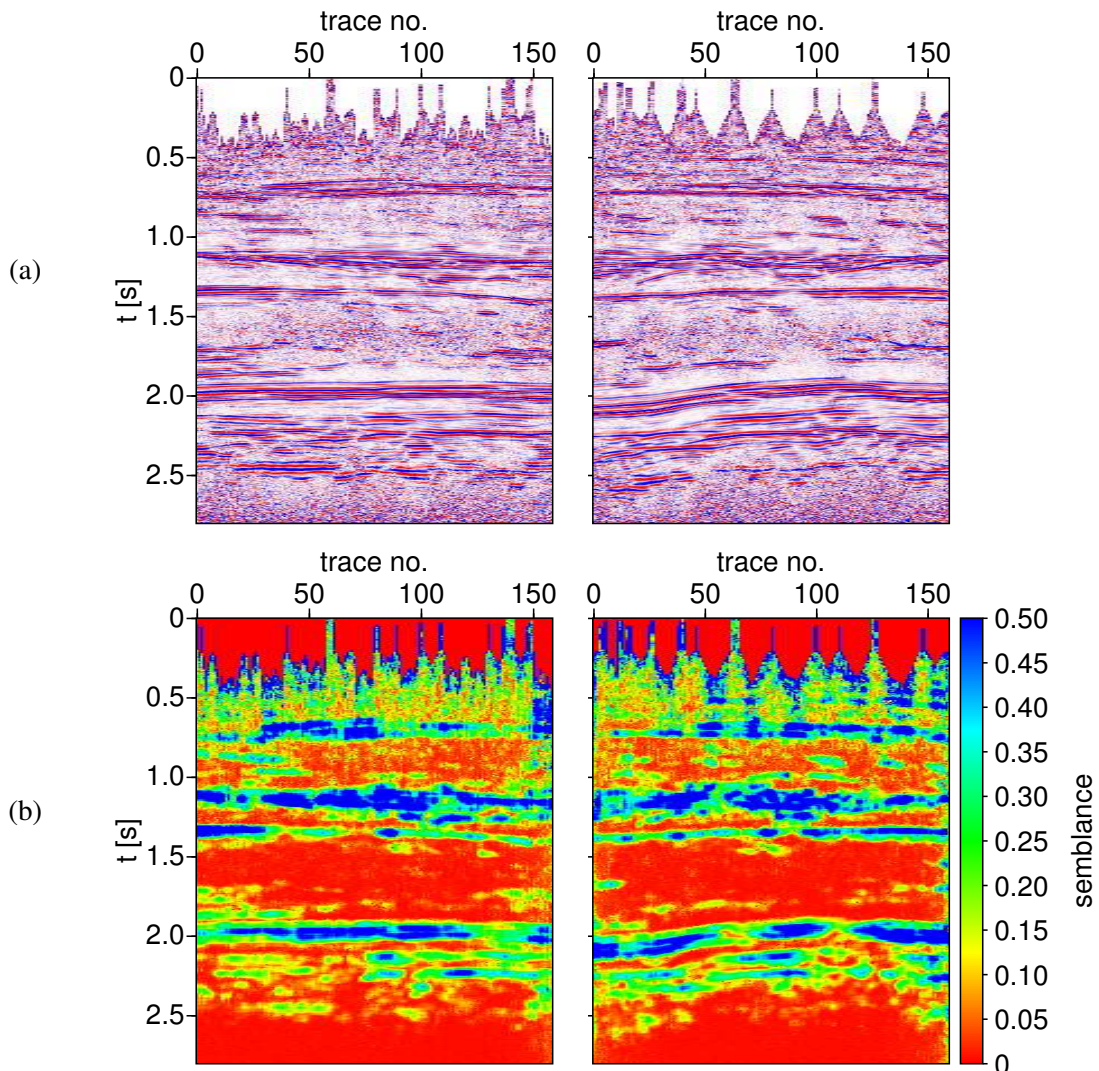


Figure 6.10: CRS stack results for profile 1 (left column) and profile 2 (right column): (a) stacked section, (b) corresponding coherence section. All sections are obtained with smoothed stacking parameters.

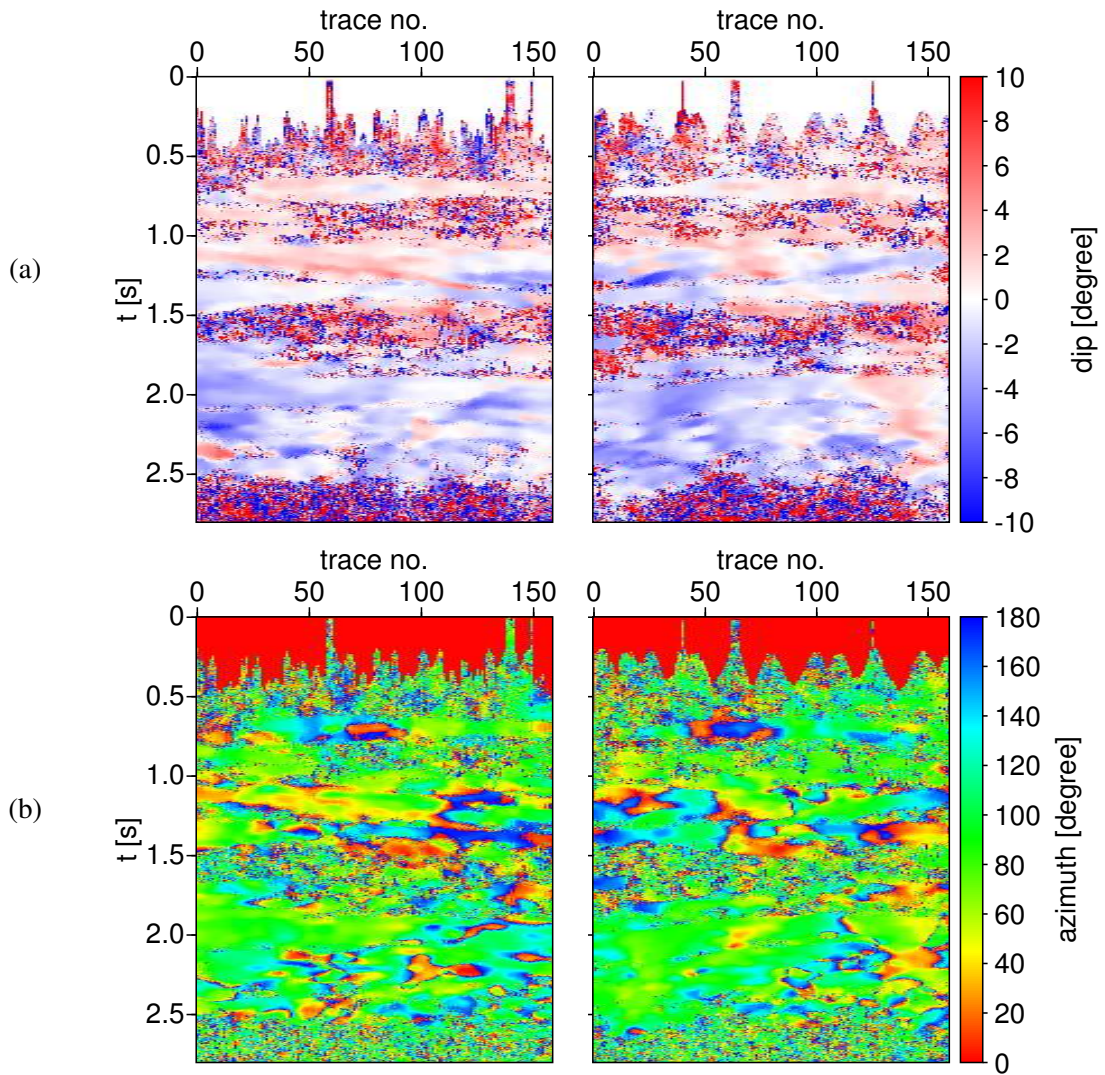


Figure 6.11: CRS stack results for profile 1 (left column) and profile 2 (right column): (a) normal ray emergence dip, (b) normal ray emergence azimuth. All stacking parameters are smoothed in an event-consistent way.

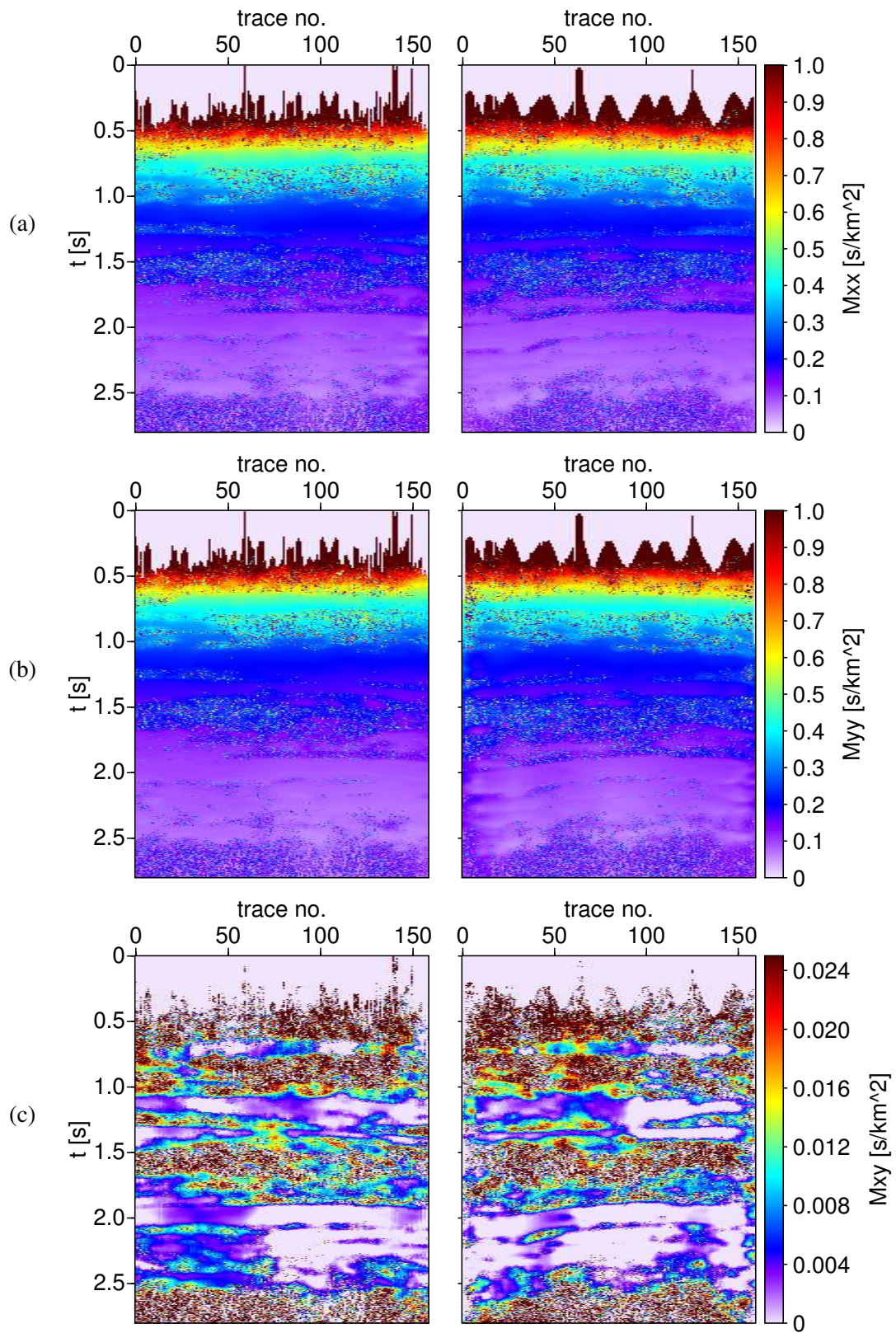


Figure 6.12: CRS stack results for profile 1 (left column) and profile 2 (right column): (a) M_{xx} , (b) M_{yy} , and (c) M_{xy} . All stacking parameters have been smoothed in an event-consistent way.

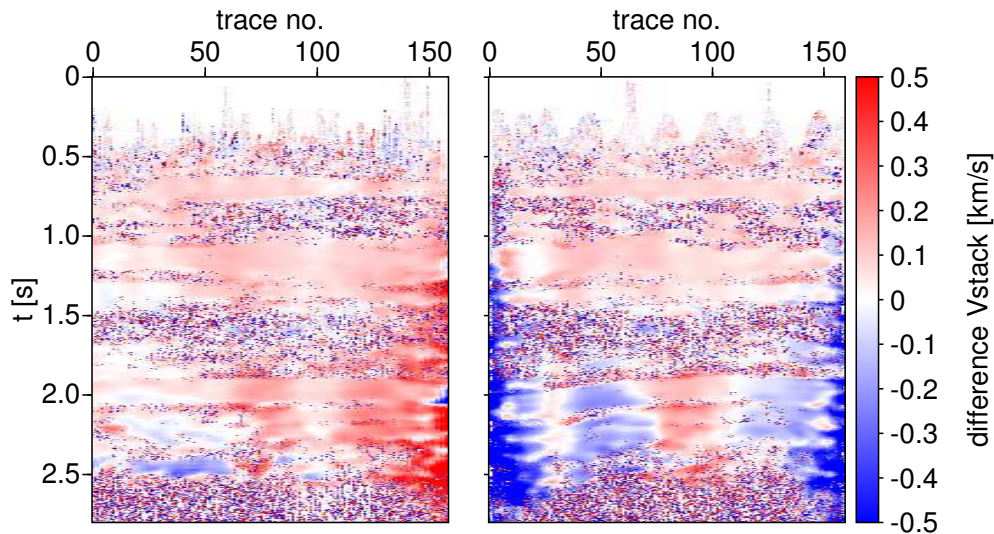


Figure 6.13: Difference in stacking velocity calculated for the x-direction and y-direction along profile 1 and 2. Note the systematic variation at the edges of the sections. See text for details.

operator is depicted at the bottom of Figure 6.10. The semblance reaches values higher than the cut-off value of 0.5 used for display purposes. Generally, the same prominent reflection events as in the stack sections can be identified. The semblance values are in many parts a little bit lower than before the application of the event-consistent smoothing (the corresponding volume is not shown here) because the operator with smoothed stacking parameters is no longer a best fit operator. However, as unphysical fluctuations have been removed from the stacking parameters, coherent events after the smoothing are considered to deliver more reliable input for the tomographic inversion.

Figure 6.11 shows the emergence direction of the normal ray. At the top the dip angle, i. e., the angle between the normal ray and the vertical, is shown. The corresponding azimuthal direction is shown at the bottom of Figure 6.11. The azimuth is measured counterclockwise starting at the positive x-axis. It reaches only values up to 180° as the emergence angle is signed. For details see Müller (2003). The dip angle is generally small, hardly exceeding 10° . The azimuth covers the entire range.

Figure 6.12 shows the three independent elements of matrix \mathbf{M} defined in equation (3.2c). The elements on the main diagonal show the characteristic behaviour of decreasing values with increasing traveltimes. The difference between the two elements is small, indicating a small azimuthal dependence of the second traveltimes derivatives. It will be demonstrated in the tomographic inversion that this difference is significant. The off-diagonal element demonstrates this azimuthal dependence as the values vary along the reflection events.

In Figure 6.13 the difference between stacking velocity in y-direction and x-direction calculated according to equation (3.6) is depicted. Although basically showing the same information as a comparison between the corresponding sections of second traveltimes derivatives, the azimuth-dependence of second-order traveltimes information is more obvious in this kind of depiction. It translates to only a few samples moveout difference within the stacking aperture. However, stacking velocity is an integral quantity. In the next subsection, I demonstrate that this azimuth-dependence may have a large influence when it is translated to an interval velocity model. To which amount these effects are due to the improper handling of surface topography is difficult to estimate and not investigated here.

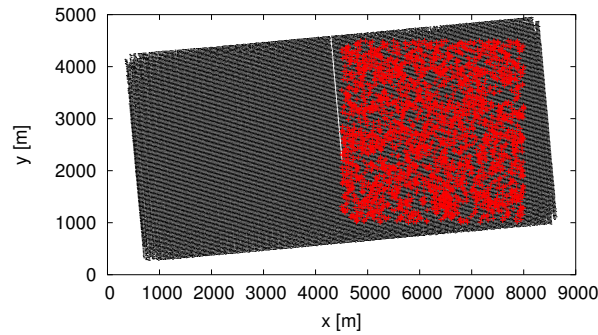


Figure 6.14: Pick locations. ZO locations are depicted with black dots, pick locations with red crosses.

A striking effect is the systematic variation of stacking velocity at the edges of the sections. Whereas the stacking velocity difference is positive at the right-hand edge of profile 1 it is negative at the edges of profile 2. I interpret this as a systematic effect due to the different azimuthal distribution of maximum offsets, especially at the edges of the prestack data volume. This yields a varying influence of the spread-length bias on the final stacking parameter values.

6.2.3 Tomography

The CRS results depicted above served as input for the tomographic inversion scheme described in Chapter 3. Two inversions have been performed: one inversion using only second traveltime derivatives for azimuth $\phi = 0^\circ$, i. e. only using the matrix element M_{xx} , and a second inversion using the complete matrix of second traveltime derivatives. These inversions are referred to as single-azimuth and multi-azimuth inversion. The input has been extracted from the CRS results using the strategy presented in Section 3.2.5. Both inversions have been performed with data from the same pick locations.

The input was picked on the ZO traces marked with red crosses in Figure 6.14. After removing some outliers and reducing the remaining number of picks, 2504 picks have been used as input to the inversions. Outliers are characterised by unreasonable stacking velocity values. Figure 6.15 shows all picks used in the inversions, depicted as stacking velocity versus ZO one-way traveltime. Stacking velocity has been calculated according to equation (3.6). The small azimuthal dependence of stacking velocity is obvious in this figure.

Both inversions were set up with common parameters summarised in Table 6.1. The velocity model was parametrised with a total number of 1216 B-spline coefficients: 8 nodes in x-direction covering the range from 4500 m to 8000 m, 8 nodes in y-direction covering the range between 1000 m and 4500 m, and 19 nodes in the vertical direction from the surface at 0 m up to 4500 m. The spacing was 500 m in both horizontal dimensions. In the vertical direction, the node spacing was 250 m. The initial model was chosen as a simple vertical gradient model, starting with 2000 m/s at the surface and a vertical gradient of 0.7 s^{-1} .

Constant weights for each type of input data have been used: traveltimes, their second derivatives, and the spatial locations had a weight of $\sigma_T = \sigma_M = \sigma_x = 1$ whereas the horizontal slowness components were weighted with $\sigma_p = 2$. No individual weighting of each pick using data-derived uncertainties was applied.

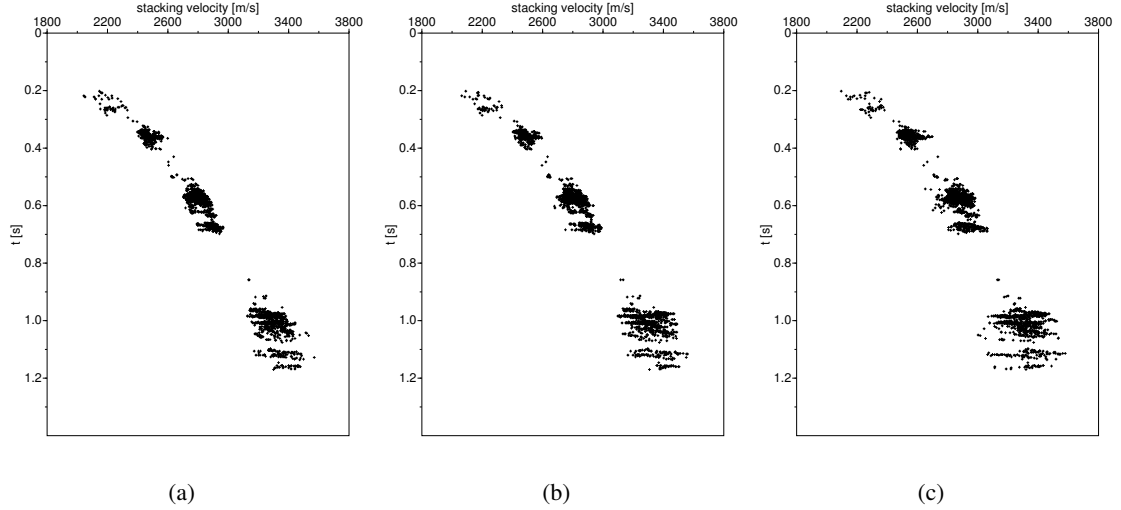


Figure 6.15: Picks depicted as stacking velocity versus ZO one-way time for three azimuthal directions: (a) x-direction, (b) xy-direction, and (c) y-direction.

Nodes		Inversion	
n_x	8	Runge-Kutta step size	30 m
n_y	8	no. of iterations	15, 10
n_z	8	max. condition number	5000
Δx	500 m	σ_T	1.0
Δy	500 m	$\sigma_{p_x}, \sigma_{p_y}$	2.0
Δz	250 m	σ_x, σ_y	1.0
extent in x	$4500 \text{ m} \leq x \leq 8000 \text{ m}$	$\sigma_{M_{xx}}, \sigma_{M_{xy}}, \sigma_{M_{yy}}$	1.0
extent in y	$1000 \text{ m} \leq y \leq 4500 \text{ m}$	σ_v	50.0
extent in z	$0 \text{ m} \leq z \leq 4500 \text{ m}$	σ_{v_q}	50.0
Initial model		Regularisation	
v_0	2000 m/s	$\epsilon_{xx}, \epsilon_{yy}, \epsilon_{zz}$	0.001
$\partial v / \partial z$	0.7 s^{-1}	ϵ	10^{-12}

Table 6.1: Parameterisation of the tomographic inversions.

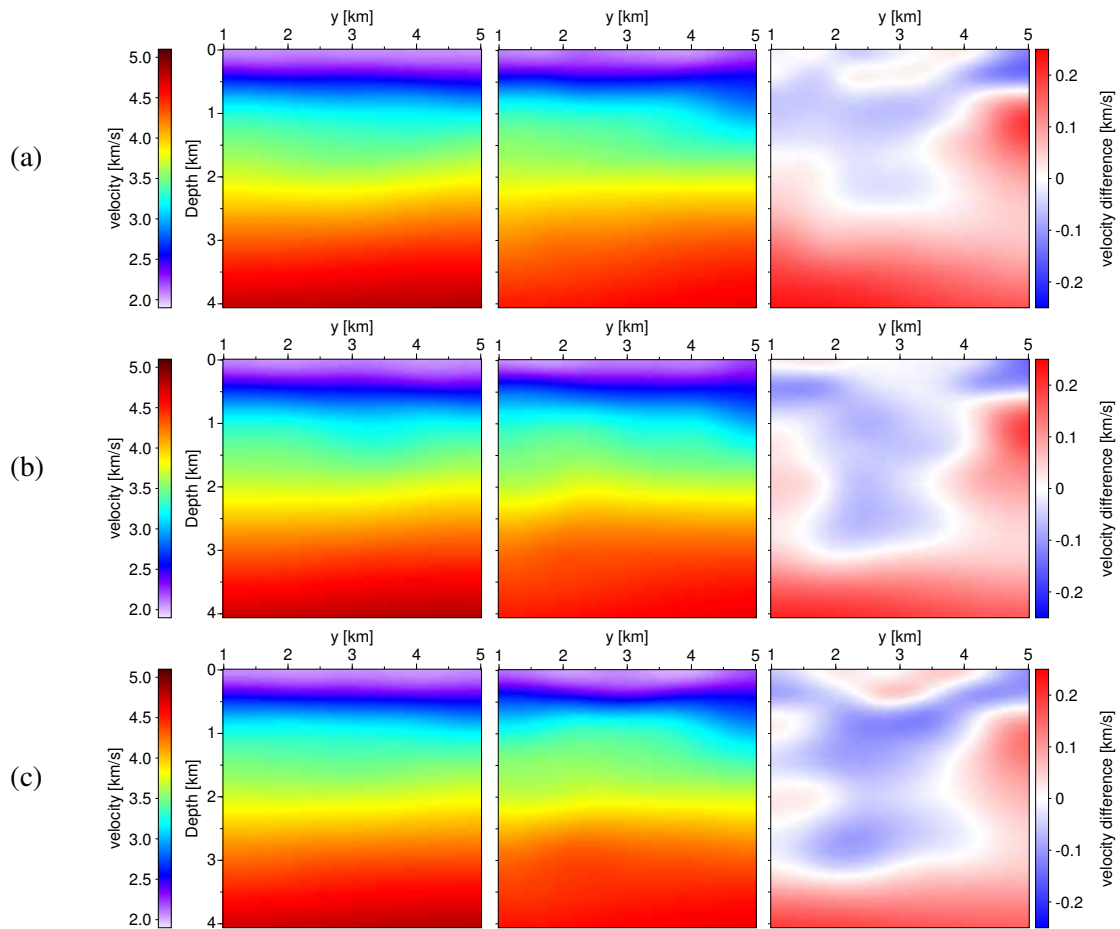


Figure 6.16: Slices through the velocity models obtained with the single azimuth (left) and the multi-azimuth (centre) inversion at (a) $x = 5000$ m, (b) $x = 6000$ m, and (c) $x = 7000$ m. The right-hand column shows the corresponding difference.

The final velocity model was constrained to a value of approximately 2050 m/s at the surface. This constraint was less weighted than the picked input data using $\sigma_v = 50$. Additionally, the velocity gradient has been forced to follow the initial normal ray direction at the NIP-locations. The according weight σ_{vq} was set to 50.

The regularisation weights for the minimisation of the second spatial derivatives of the velocity model were set to 0.001, whereas the value for ϵ was set to 10^{-12} in order not to minimise the velocity itself.

Ray tracing was performed using a second-order Runge-Kutta scheme with a vertical step size of 30 m. The single-azimuth inversion was performed with 15 iterations, the multi-azimuth inversion stopped after 10 iterations because the cost-function was no longer decreasing.

A comparison between both final velocity models is shown along several cross sections through the model volume in Figure 6.16. The difference between both models reaches values up to 250 m/s at the edges of the sections and values up to 150 m/s in the centre. This results in significant differences in the final images after depth migration. This demonstrates the effect of azimuth-dependent second-order traveltme information on the interval velocity model in the depth domain.

A first indicator of the consistence between the velocity model and the prestack data is the distribution

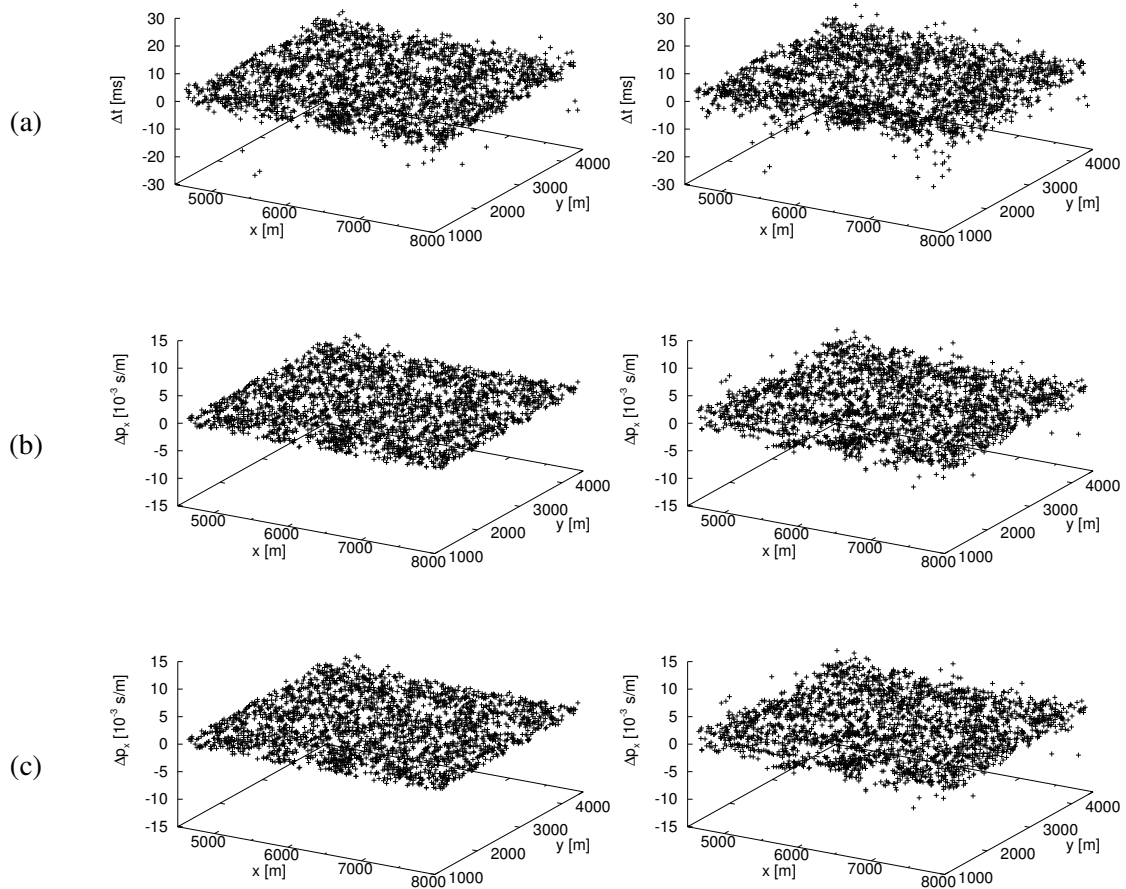


Figure 6.17: Final residual errors as a function of pick location for the single-azimuth (left) and the multi-azimuth (right) inversion: (a) one-way traveltime t , (b) x-component, and (c) y-component of the horizontal slowness.

of final residual errors in the forward-modelled attribute values. Figures 6.17 and 6.18 show these errors for the ZO one-way traveltime T , the two horizontal slowness components, and the three independent elements of matrix \mathbf{M} . The corresponding errors of the normal ray emergence location are not shown here, because they are below 3 m in both inversions. Compared to the mean CMP spacing of 25 m this is an almost perfect match.

The residual errors of the normal ray traveltime are generally in the range of 20 ms, i. e., at most five samples in the poststack volume. Thus, together with the small lateral error, the pick locations are well recovered. The errors in the horizontal slowness components are very small, too. This indicates that the time dip at the pick locations has been accurately recovered. Generally, the residual errors show slightly larger fluctuations in the multi-azimuth inversion.

Similar observations are made for the errors in the elements of the second traveltime derivative matrix. In the single-azimuth inversion, the used element M_{xx} is reconstructed very well. The same holds for the multi-azimuth inversion. However, here the residual error shows a little bit larger fluctuations. The other two elements of matrix \mathbf{M} have been recovered well in the case of the multi-azimuth inversion. These elements have been forward modelled in the final model of the single-azimuth inversion using the NIP-locations and initial normal ray directions obtained in that inversion. In the single-azimuth

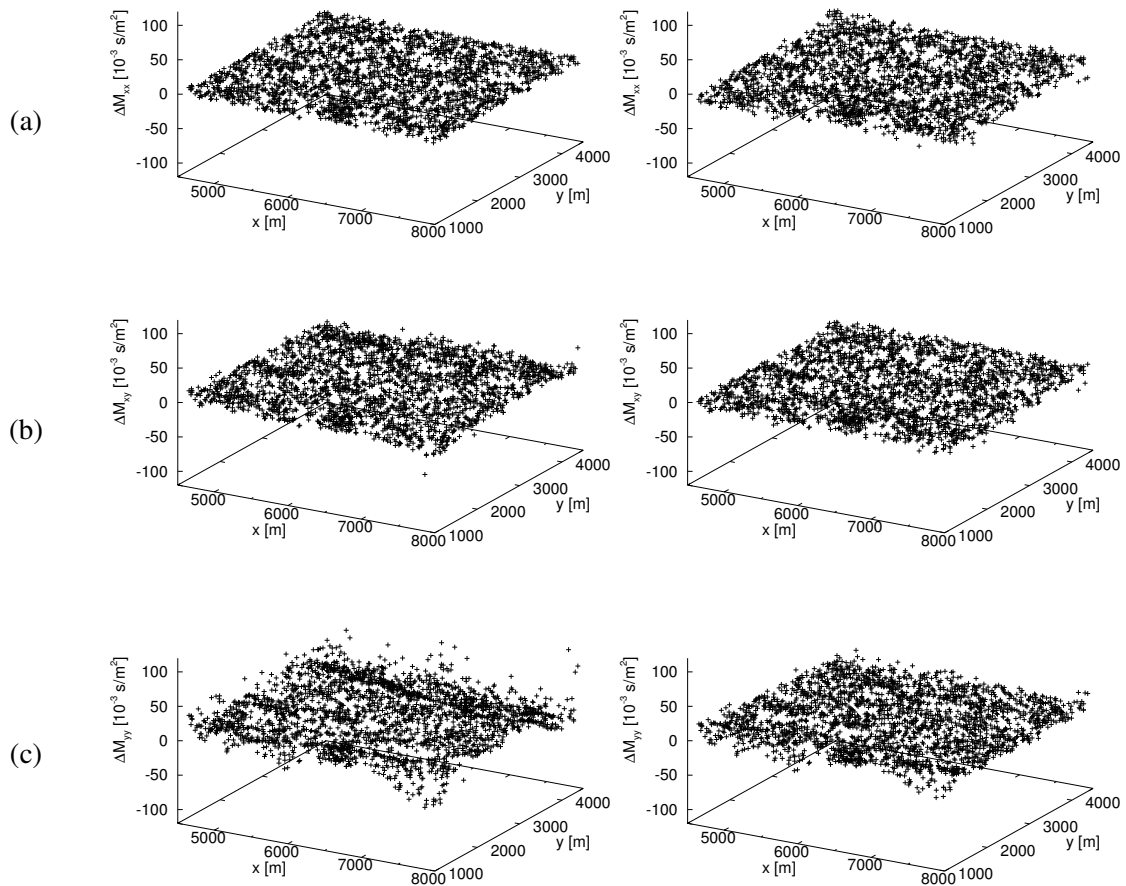


Figure 6.18: Final residual errors as a function of pick location for the single-azimuth (left) and multi-azimuth (right) inversion: (a) M_{xx} , (b) M_{xy} , and (c) M_{yy} .

inversion, especially the element M_{yy} shows significant residual errors. This clearly indicates that the model obtained by single-azimuth inversion is not consistent with the prestack data which show azimuthal dependence of second-order traveltime information.

The spatial distribution of NIP-locations is visualised in Figure 6.19: the distances between corresponding NIP-locations of both inversions reach values up to 100 m and, for a few selected picks, up to 180 m. This means that the same point on a reflector is mapped 100 m differently using the velocity models obtained in both inversions for depth migration. This has a large impact on any structural interpretation of the final depth migrated image.

6.2.4 Migration

The final quality indicator presented here are selected results of Kirchhoff depth migration obtained with the velocity models of both inversions. For displaying purposes, a gain function has been applied to all migrated sections. Figures 6.21 and 6.22 show CIGs in an azimuth-dependent way. For each output location, six CIGs are obtained each using only offset vectors within a specific azimuth sector of 60° . The used azimuth sectors are depicted in Figure 6.20. Offsets up to 4000 m have been used in the migration. The numbering in Figure 6.20 corresponds to the order of displaying the CIGs for one

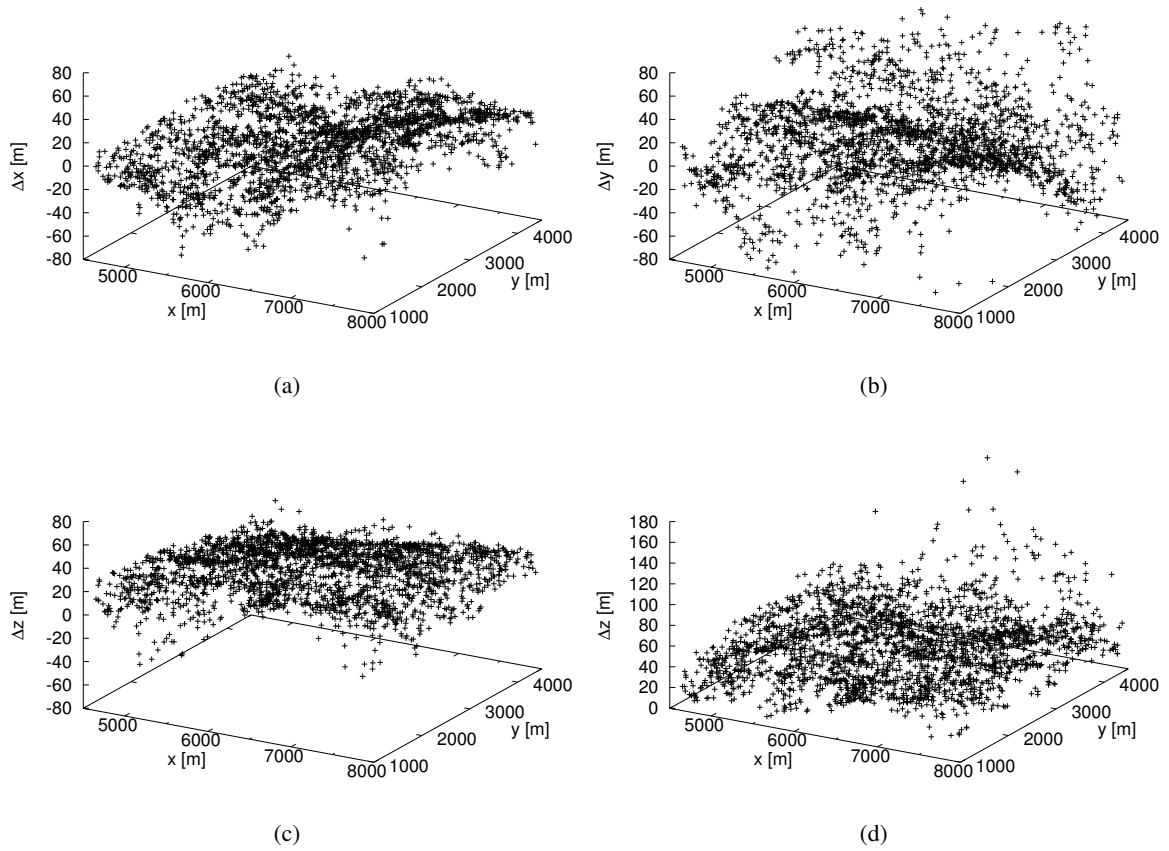


Figure 6.19: Lateral distance between the NIP-locations obtained with single-azimuth and multi-azimuth inversion as a function of pick location: (a) x-component, (b) y-component, (c) z-component, and (d) absolute displacement.

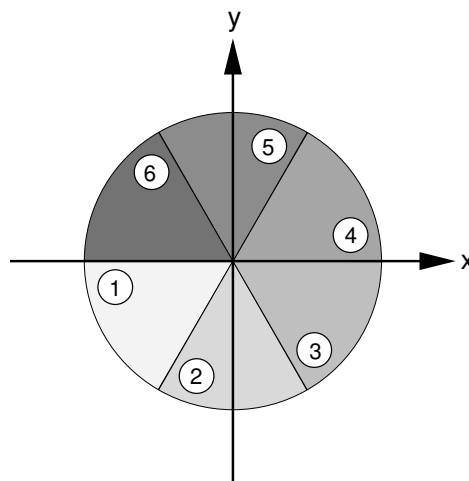


Figure 6.20: Definition of offset azimuth bins. Bins are numbered counterclockwise starting from the negative x-axis. Each bin has a size of 60°.

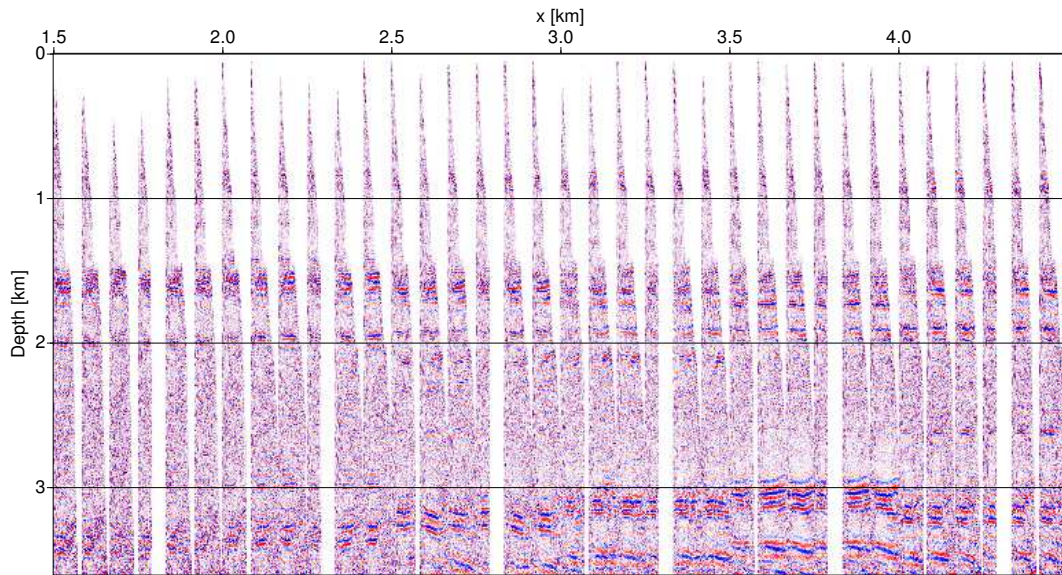
output location. CIGs with azimuths nearly parallel to the x-axis include only smaller offsets because the maximum offset found in the prestack data in these directions is generally smaller. This can be seen in Figure 6.8.

The residual moveout of coherent events in the CIGs shows no clear azimuth-dependence. This is a result of the lateral and vertical velocity variations which lead to a complex, model-based transformation of the residual errors of the elements of matrix \mathbf{M} into residual moveout. The residual moveout in the CIGs computed with the multi-azimuth velocity model is slightly larger for shallow events but for almost all deep events it is significantly smaller than in the CIGs for the single-azimuth inversion.

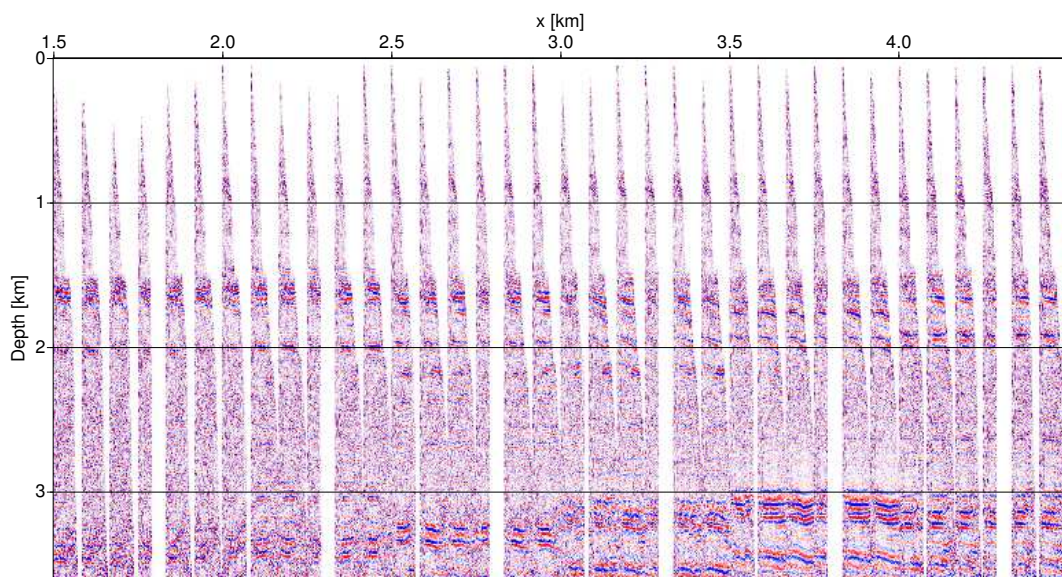
Figure 6.23 shows slices through the poststack depth migrated CRS stack volume. The structural differences between the results obtained with the two velocity models are best seen in the sections in the left-hand column for the reflectors between $z = 3000$ m and $z = 4000$ m. The differences between both velocity models lead to reflector dip changes in this region.

6.2.5 Conclusions

I have demonstrated the application of the tomographic inversion scheme based on second-order approximations of diffraction traveltimes to a 3D real data set. I compared an inversion using second-order traveltimes information in just a single azimuth with an inversion using the complete matrix of second traveltimes derivatives. The presented results demonstrate the applicability of the proposed inversion scheme. Furthermore, they show the importance of azimuth-dependent analysis of second-order traveltimes information and its incorporation in the inversion scheme.

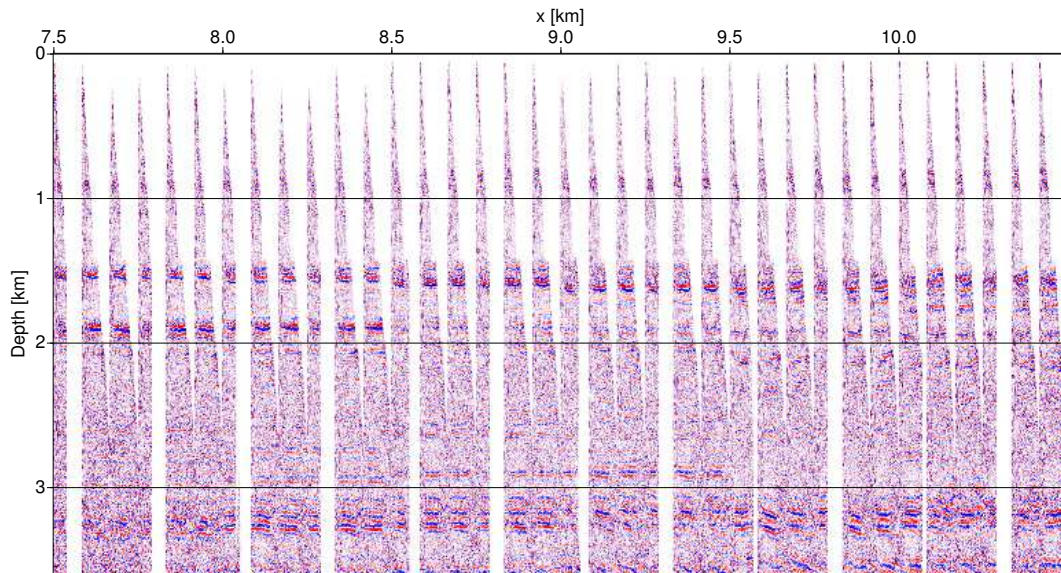


(a)

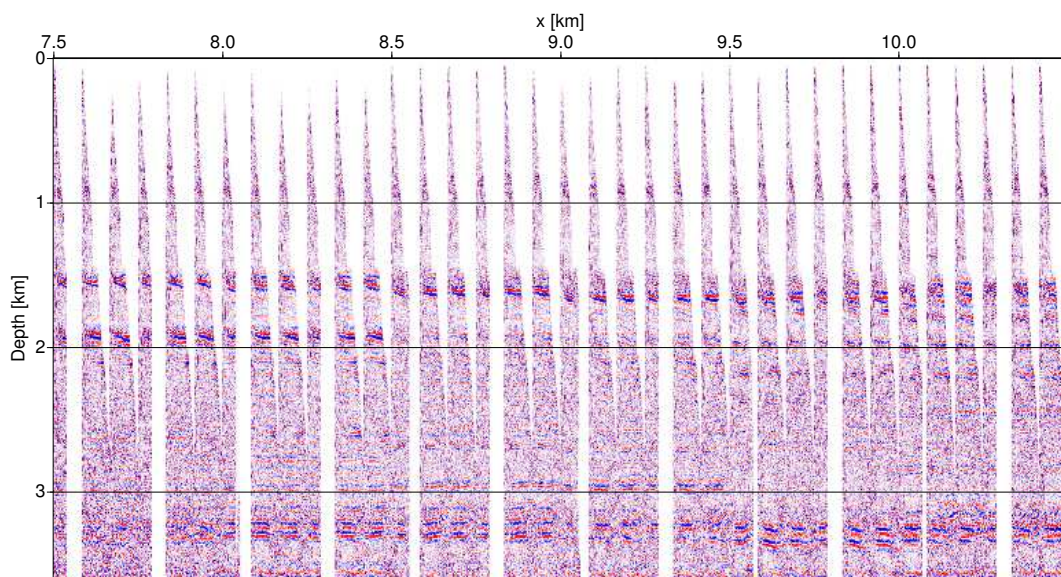


(b)

Figure 6.21: Azimuthally dependent CIGs at $x = 6500$ m obtained with the velocity models from a) the single-azimuth and b) the multi-azimuth inversion.



(a)



(b)

Figure 6.22: Azimuthally dependent CIGs at $y = 2500$ m obtained with the velocity models from the a) single-azimuth and b) multi-azimuth inversion.

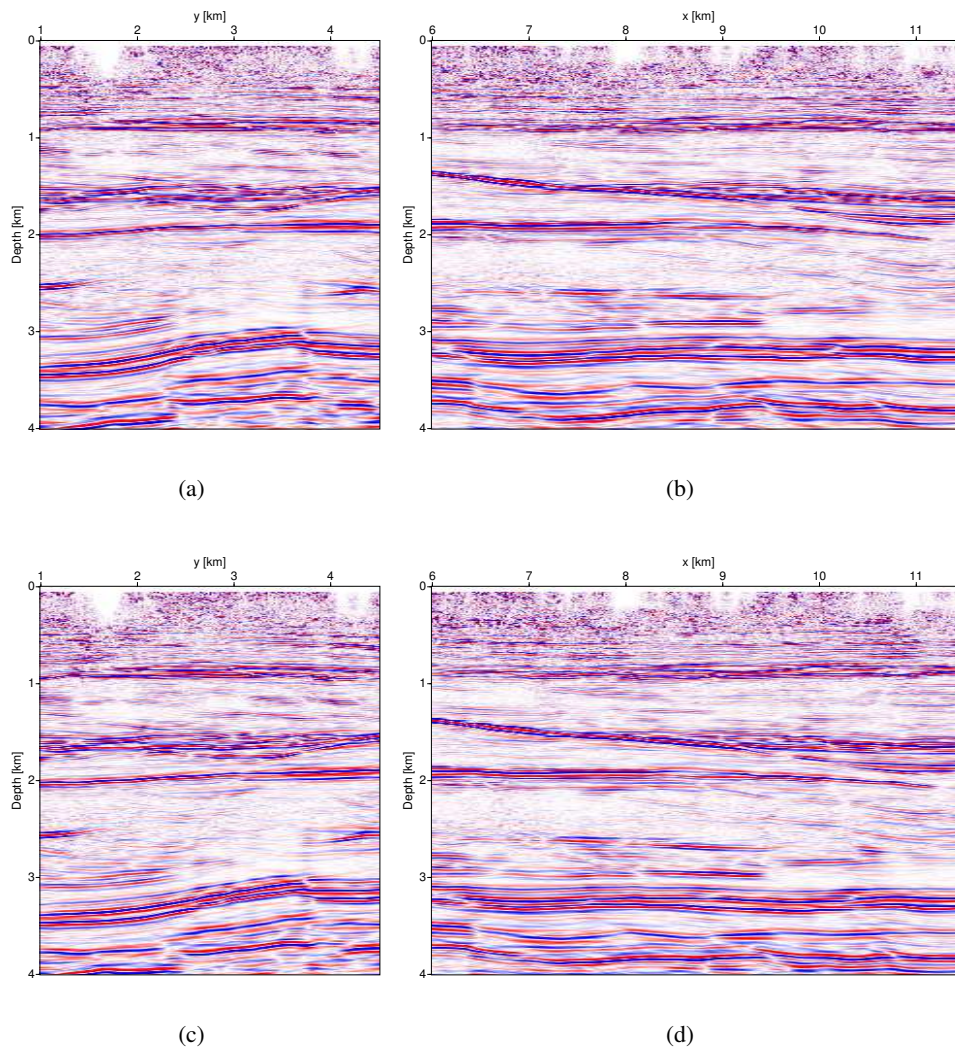


Figure 6.23: Depth-migrated CRS stack sections using the velocity model from the single-azimuth (top) and multi-azimuth (bottom) inversions: (a) and (c) at $x = 6500$ m, (b) and (d) at $y = 2500$ m.

Chapter 7

Summary and outlook

In this thesis, I have discussed two tomographic inversion schemes for the construction of smooth isotropic velocity models suitable for depth migration of seismic data. Both methods make use of diffraction traveltimes, i. e., of the kinematic responses of diffractors/point sources in depth. Each input data point of both inversion schemes is associated with such a point source. For this reason, each data point can be treated independently of all others and no continuous reflector interfaces need to be included in the model description. For each data point an isolated reflection point and the local reflector dip at that position is part of the model description. The smooth velocity model is parameterised in terms of B-splines. The input data for both inversion methods have some quantities in common: each input data point contains attributes parameterising a zero-offset stationary point, i. e., a point of tangency between the diffraction traveltimes operator and the actual reflection event in the zero-offset stack section/volume. The link between the zero-offset time domain and the depth domain is established by the concept of the normal ray. The attributes characterising a zero-offset stationary point are its lateral position, the zero-offset one-way traveltimes along the normal ray and the horizontal slowness of the emerging normal ray. The identification of these points in a zero-offset stack section/volume with a considerably increased signal-to-noise ratio is much easier than picking input information directly in the seismic prestack data as it is, for example, necessary in classical traveltimes tomography. As each data point is treated independently of all others there is no need to identify continuous reflection events over several traces, the identification of locally coherent events is sufficient. Furthermore, this identification of locally coherent events may be highly automated.

The first inversion method has already been extensively discussed by [Duvencq \(2004a\)](#). I have called this method NIP-wave tomography. In this thesis, I have made several extensions to that method and the associated data processing chain based on the common-reflection-surface (CRS) stack method. Beside the above mentioned parameterisation of zero-offset stationary points, the second traveltimes derivative at the normal ray emergence location with respect to the source-receiver distance are input to NIP-wave tomography. These second traveltimes derivatives, matrices in the 3D case and scalar quantities in the 2D case, can be used within the validity range of the NIP-wave theorem to parameterise second-order approximation to the true diffraction traveltimes functions. Alternatively, it can be physically interpreted as the second traveltimes derivative associated with the NIP-wave which is emitted from a point source placed in depth.

All information needed in NIP-wave tomography is automatically extracted from the seismic prestack data by means of the CRS stack method. I have presented strategies to precondition the input data and to determine its uncertainty in a data-driven way. These strategies use a small window/volume aligned

with the reflection event in the CRS stacked section/volume. Inside this event-aligned window/volume all CRS stacking parameters are subject to a combination of median filtering and averaging to remove outliers and unphysical fluctuations. Alternatively, the CRS stacking parameters within the event-aligned volume may be used for a local robust linear regression which yields not only smoothed stacking parameters but also an estimate of their uncertainties. Based on the idea of the event-aligned window/volume I have presented a method to identify valid pick locations and to extract the necessary input information for NIP-wave tomography at these locations in a highly automated way from the CRS results. The method is based on the coherence estimated along the CRS stacking operator and uses additional criteria for the identification of locally coherent events which are similar to those used in the pre-conditioning.

I have applied the uncertainty estimates of the CRS stacking parameters as data-driven covariance operators in the iterative solution of the inverse problem associated with NIP-wave tomography. This inverse problem has been formulated as the minimisation of a cost-function in the least squares sense. In this way not only each type of input data is weighted as in the original implementation but each input data quantity has its individual weight. The application of data-driven covariance operators has been presented using a synthetic 2D data example.

The implementation of NIP-wave tomography has been parallelised based on the message-passing-interface (MPI). This step allowed efficient testing of the code especially in the 3D case. I have extended the 3D implementation to handle all three independent elements of the matrix of second traveltimes derivatives. This allows to include azimuth-dependent second-order traveltimes information in the inversion process. However, the extraction of these azimuth-dependent information from the seismic prestack data needs acquisition geometries with offset vectors covering the complete azimuthal range. In this thesis, I have presented a comparison of NIP-wave tomography results obtained with single-azimuth and multi-azimuth second-order traveltimes derivatives using a 3D real data example.

An important aspect covered in this thesis is to overcome the limitation to second-order traveltimes approximations inherent in NIP-wave tomography. The velocity models obtained by NIP-wave tomography have to be considered as only being consistent with the seismic prestack data up to the second order. I have presented a new tomographic inversion scheme for the 2D case which, in principle, has no limitations regarding the complexity of the subsurface velocity distribution. The method aims at minimising the kinematic difference between common-reflection-point gathers and common-image-gathers. The diffraction traveltimes function forward-modelled in the current velocity model is applied as a moveout function to the seismic prestack data. This results in a small prestack data volume for each input data point in which all traces are sorted with respect to the illumination and scattering angles defined at the reflection point. Time is measured with respect to the forward-modelled diffraction traveltimes. From these small prestack data volumes a CRP gather is extracted selecting all traces with varying scattering angles and an illumination angle coinciding with the current estimate of the reflector dip. When the velocity model is consistent with the prestack data the reflection event has to be located at time zero for all scattering angles. I have demonstrated that this criterion alone is not sufficient to ensure the consistency of the velocity model with the seismic prestack data when isolated, sparsely selected depth points are considered. It has further to be ensured that the signal at time zero in a CRP gather actually is due to a reflection at the considered depth point. This implies that the forward-modelled diffraction traveltimes operator is tangent to the prestack data for all scattering angles. This criterion can be checked considering common-scattering-angle sections extracted from the small moveout-corrected prestack data volumes. The time dip of the reflection event has to vanish in these sections. The second traveltimes derivatives used as input in NIP-wave tomography are

replaced in the new inversion scheme by time and time dip residuals extracted along the model-based CRP trajectory for each input data point. Together with these time and time dip residuals the misfit between the input and forward-modelled zero-offset stationary point descriptions is minimised. The inversion scheme requires much more human interaction as the time and time dip residuals have to be picked again in each iteration of the inversion and the picking results have to be carefully examined to remove outliers. The time dip residuals in the angle domain are related to slowness misfits at the source and receiver locations by model-dependent transformation factors. This relationship has been derived and applied in the implementation of the method. The inversion scheme is presented as a stand alone application using a 2D synthetic data example.

As the new inversion method is implemented using the same model description and partially also the same zero-offset input data, it is an ideal successor of NIP-wave tomography in a processing chain. This allows to set up a complete velocity model building work flow starting with the CRS stack followed by NIP-wave tomography and optionally finalised by the new inversion method used as an updating technique. This velocity model building work flow has been demonstrated using a complex 2D synthetic data example.

As a future task the inversion method using model-based diffraction traveltime functions may be extended to the 3D case. For that purpose, efficient strategies to extract time and time dip residuals in an azimuth-dependent way from the, in the 3D case, five-dimensional small moveout-corrected prestack data volumes have to be developed. This would allow to extend the 2D velocity model building work flow presented in this thesis to the 3D case.

A serious problem in the automated extraction of input data for NIP-wave tomography is the possible picking of locally coherent events related to multiply reflected energy. Efficient criteria for the identification of such picks need to be developed.

An important practical aspect that needs further work is the development of an automatic regularisation strategy which better accounts for regions less constrained by the input data. This can be done based on the current implementations adjusting the way of minimising the curvature of the velocity model. Different strategies appear with a change of the domain in which reliable pick locations are identified. A regular distribution of picks in the zero-offset stacked section/volume usually leads to clustering of the associated reflection points in depth. This results in an irregular coverage of the depth model by rays and may lead to instabilities in the inversion process. Picking in a time-migrated section or even in a poststack depth migrated section would yield a more regular ray coverage in the depth domain.

Appendix A

Velocity model in terms of B-splines

The velocity model description used in the inversion methods is based on B-splines (de Boor, 1978). In this appendix a short summary of the used velocity model parameterisation is given. A more detailed treatment can be found in Duveneck (2004a).

B-splines of degree m allow to represent a spline function f of the same degree as a weighted sum of defined values at a sequence of node locations. Let $\beta_i(x)$ be the B-spline basis function and f_i the defined value at the i th node point. The B-spline representation of function f then reads

$$f(x) = \sum_i^{n_x} f_i \beta_i(x). \quad (\text{A.1})$$

For the velocity model description the sequence of node points is assumed to be strictly increasing:

$$\dots < x_{i-1} < x_i < x_{i+1} < \dots$$

The node intervals do not necessarily need to be of constant length. B-splines of degree m only are non-zero on $m + 1$ consecutive node intervals and may be defined recursively by

$$\beta_{[x_i, \dots, x_{i+m+1}]}^{[m]}(x) = \frac{x - x_i}{x_{i+m} - x_i} \beta_{[x_i, \dots, x_{i+m}]}^{[m-1]}(x) + \frac{x_{i+m+1} - x}{x_{i+m+1} - x_{i+1}} \beta_{[x_{i+1}, \dots, x_{i+m+1}]}^{[m-1]}(x). \quad (\text{A.2})$$

The B-spline basis function of degree zero takes the values

$$\beta_{[x_i, x_{i+1}]}^{[0]} = \begin{cases} 1 & \text{for } x \in [x_i, x_{i+1}) \\ 0 & \text{else} \end{cases}. \quad (\text{A.3})$$

In NIP-wave tomography presented in chapter 3 continuous third derivatives of the velocity are needed in the computation of the Fréchet derivatives for the second traveltime derivatives. Therefore, a representation of the velocity model with B-splines of degree $m = 4$ is chosen. In the inversion scheme using model-based diffraction traveltime operators only continuous second derivatives of the velocity model are required. In principle, a representation of the velocity model with B-splines of degree $m = 3$ could be chosen for that reason. This would allow even more locality in the velocity model description. However, to enable the usage of this inversion scheme as an updating technique for the models determined with NIP-wave tomography the velocity model is represented with B-splines of degree $m = 4$ as well. All equations in the following are given for the 3D case. The axes of the global

Cartesian coordinate system are denoted by x , y , and z . The corresponding expressions for the 2D case are obtained omitting the y -dimension.

Part of the model description in the 3D inversion scheme are three user-defined node sequences with n_x , n_y , and n_z nodes, respectively. These nodes define the grid of B-spline coefficients. For the representation of the velocity model at least $m + 1$ additional nodes are required in each spatial dimension. The location of these node points is automatically determined with increasing node intervals to the margins of the model. The values of the outermost user-defined B-spline coefficients are copied for these additional node points. This ensures that the velocity is extended virtually constant in all directions and has reasonable values even outside the region which is constrained by the data.

The implementation of the presented inversion schemes uses Cartesian coordinates with the positive z -axis pointing upwards. All depth locations in the inversion have negative z -values. However, the node sequence for the z -direction is specified with depth increasing with increasing z -value. Therefore, $-z$ is used for velocity calculations and all velocity derivatives with respect to the z -coordinate are multiplied by $(-1)^q$ where q denotes the degree of the derivative with respect to the z -coordinate.

The 3D velocity model representation with the B-spline coefficients v_{jkl} reads

$$v(x, y, z) = \sum_{j=1}^{n_x} \sum_{k=1}^{n_y} \sum_{l=1}^{n_z} v_{jkl} \beta_j(x) \beta_k(y) \beta_l(-z) . \quad (\text{A.4})$$

Its derivatives are calculated according to

$$\frac{\partial^{o+p+q} v(x, y, z)}{\partial x^o \partial y^p \partial z^q} = \sum_{j=1}^{n_x} \sum_{k=1}^{n_y} \sum_{l=1}^{n_z} v_{jkl} (-1)^q \frac{\partial \beta_j(x)}{\partial x^o} \frac{\partial \beta_k(y)}{\partial y^p} \frac{\partial \beta_l(-z)}{\partial z^q} . \quad (\text{A.5})$$

For the calculation of the Fréchet derivatives in the inversion schemes the partial derivative of (A.4) with respect to the B-spline coefficients v_{jkl} is needed. For the general 3D case these derivatives read

$$\frac{\partial v(x, y, z)}{\partial v_{jkl}} = \beta_j(x) \beta_k(y) \beta_l(-z) , \quad (\text{A.6})$$

with $j = 1, \dots, n_x$, $k = 1, \dots, n_y$, and $l = 1, \dots, n_z$.

Appendix B

Calculation of Fréchet derivatives

The calculation of the Fréchet derivatives needed in the presented inversion schemes is based on ray and ray perturbation theory as it is described in Chapter 2. Firstly, the Fréchet derivatives for 3D NIP-wave tomography are outlined. They are based on the formulation of ray and ray perturbation theory with the reduced Hamiltonian approach in global Cartesian coordinates. Then, the necessary expressions for the 2D case of NIP-wave tomography and the 2D inversion using model-based diffraction traveltimes operators are derived. The formulation in these cases is based on ray-centred coordinates. The Fréchet derivatives are obtained considering infinitesimal small perturbations in the following equations, i. e., implicitly considering the limit

$$\lim_{\Delta \mathbf{m} \rightarrow 0} \frac{\Delta \mathbf{d}}{\Delta \mathbf{m}}.$$

B.1 Fréchet derivatives in global Cartesian coordinates

In 3D NIP-wave tomography with azimuth-dependent second traveltimes derivatives the Fréchet derivatives have formally be written in Section 3.2.6 as

$$\frac{\partial (T_i, \mathbf{p}_i, \mathbf{M}_i, \mathbf{x}_{m_i})}{\partial (x_i, y_i, z_i, e_{x_i}, e_{y_i}, v_{jkl})}, \quad (\text{B.1})$$

where $i = 1, \dots, n_{\text{data}}$, $j = 1, \dots, n_x$, $k = 1, \dots, n_y$, and $l = 1, \dots, n_z$. The analogous expression for the inversion with narrow azimuth second-order traveltimes information is obtained replacing the matrix \mathbf{M} by M_ϕ . The calculation of the Fréchet derivatives uses the formulation of ray and ray perturbation theory in reduced Cartesian coordinates introduced in Chapter 2.

To simplify the notation in this chapter, the coordinate axes are named x , y , and z instead of x_1 , x_2 , and x_3 . Quantities at the NIP location have a subscript 0 whereas quantities at the measurement surface are indicated with a subscript 1.

B.1.1 Fréchet derivatives of \mathbf{x}_m and \mathbf{p}

The Fréchet derivatives of \mathbf{x}_m and \mathbf{p} are obtained from equation (2.42) which relates changes $\Delta \tilde{\mathbf{w}}_1$ of the phase space coordinates of a ray at the surface to perturbations $\Delta \tilde{\mathbf{w}}_0$ of the corresponding quantities

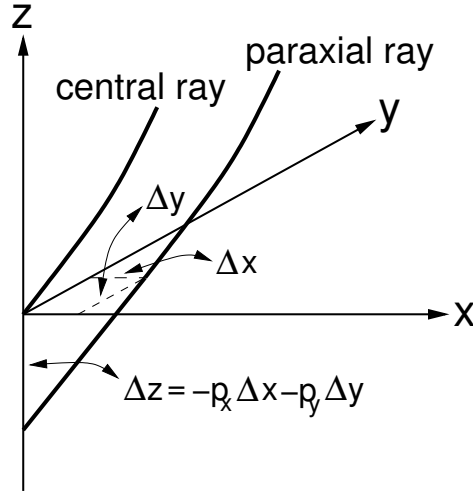


Figure B.1: Relationship between horizontal and vertical displacements at the starting position of the central ray. Without loss of generality, the origin of the global Cartesian coordinate system coincides with the starting position of the central ray.

at the NIP and perturbations of the velocity along the ray. In the here used notation equation (2.42) reads

$$\Delta \tilde{\mathbf{w}}_1 = \tilde{\mathbf{\Pi}}(z_1, z_0) \Delta \tilde{\mathbf{w}}_0 + \tilde{\mathbf{\Pi}}(z_1, z_0) \int_{z_0}^{z_1} \tilde{\mathbf{\Pi}}(z', z_0) \Delta \tilde{\mathbf{n}}(z', \Delta v) dz', \quad (\text{B.2})$$

where equation (2.43) has been applied. The integral in the above equation has to be calculated for each coefficient in the B-spline representation of the velocity model using equation (A.6) and the vector $\Delta \tilde{\mathbf{n}}$ defined in equation (2.41).

The Fréchet derivatives of \mathbf{x}_m and \mathbf{p} with respect to vertical displacements Δz_0 of the central ray's starting position are derived using the geometrical relation between horizontal and vertical displacements depicted in Figure B.1:

$$\begin{aligned} \Delta \tilde{\mathbf{w}}_1(\Delta z_0) &= \tilde{\mathbf{\Pi}}(z_1, z_0) \Delta \tilde{\mathbf{w}}_0(\Delta z_0) \\ &= \tilde{\mathbf{\Pi}}(z_1, z_0) \begin{pmatrix} -\frac{p_{x_0}}{p_{z_0}} \Delta z_0 \\ -\frac{p_{y_0}}{p_{z_0}} \Delta z_0 \\ 0 \\ 0 \end{pmatrix}. \end{aligned} \quad (\text{B.3})$$

B.1.2 Fréchet derivatives of \mathbf{M}

The matrix \mathbf{M} of second travelttime derivatives due to a point source placed at the NIP is calculated from elements of the ray propagator matrix $\tilde{\mathbf{\Pi}}(z_1, z_0)$ according to equation (2.34):

$$\mathbf{M} = \mathbf{P}_2 \mathbf{Q}_2^{-1}. \quad (\text{B.4})$$

Thus, perturbations of \mathbf{M} can be written in terms of perturbations of the individual sub-matrices of $\tilde{\mathbf{\Pi}}(z_1, z_0)$:

$$\Delta \mathbf{M} = \Delta \mathbf{P}_2 \mathbf{Q}_2^{-1} - \Delta \mathbf{Q}_2 \mathbf{P}_2 \mathbf{Q}_2^{-1} \mathbf{Q}_2^{-1}. \quad (\text{B.5})$$

The perturbation $\Delta\tilde{\mathbf{\Pi}}$ of the ray propagator matrix is obtained from equation (2.56). Using the chain rule, the expression for $\Delta\tilde{\mathbf{\Pi}}$ reads

$$\Delta\tilde{\mathbf{\Pi}}(z_1, z_0) = \tilde{\mathbf{\Pi}}_0(z_1, z_0) \int_{z_0}^{z_1} \tilde{\mathbf{\Pi}}_0^{-1}(z', z_0) [\Delta\tilde{\mathbf{S}}(z')] \tilde{\mathbf{\Pi}}_0(z', z_0) dz' . \quad (\text{B.6})$$

This may be reformulated (Duvenceck, 2004a) to

$$\Delta\tilde{\mathbf{\Pi}} = \begin{pmatrix} \mathbf{Q}_1 & \mathbf{Q}_2 \\ \mathbf{P}_1 & \mathbf{P}_2 \end{pmatrix} \begin{pmatrix} \mathbf{A}^{(1)} & \mathbf{A}^{(2)} \\ \mathbf{A}^{(3)} & \mathbf{A}^{(4)} \end{pmatrix} . \quad (\text{B.7})$$

Using the symplectic property, the expression for $\Delta\mathbf{M}$ reduces to

$$\Delta\mathbf{M} = -(\mathbf{Q}_2^{-1})^T \mathbf{A}^{(2)} \mathbf{Q}_2^{-1} . \quad (\text{B.8})$$

This means that only the upper left 2×2 sub-matrix of the integral expressions needs actually to be evaluated. Let me write the two parts of $\Delta\tilde{\mathbf{S}}$ in equation (2.50) in the following form:

$$\Delta\tilde{\mathbf{S}}_1(\Delta\tilde{\mathbf{w}}) = \begin{pmatrix} \Delta\mathbf{S}_1^{(1)}(\Delta\tilde{\mathbf{w}}) & \Delta\mathbf{S}_1^{(2)}(\Delta\tilde{\mathbf{w}}) \\ \Delta\mathbf{S}_1^{(3)}(\Delta\tilde{\mathbf{w}}) & \Delta\mathbf{S}_1^{(4)}(\Delta\tilde{\mathbf{w}}) \end{pmatrix} \quad \text{and} \quad \Delta\tilde{\mathbf{S}}_2(\Delta v) = \begin{pmatrix} \Delta\mathbf{S}_2^{(1)}(\Delta v) & \Delta\mathbf{S}_2^{(2)}(\Delta v) \\ \Delta\mathbf{S}_2^{(3)}(\Delta v) & \Delta\mathbf{S}_2^{(4)}(\Delta v) \end{pmatrix} . \quad (\text{B.9})$$

Using the relationship (2.30) between $\tilde{\mathbf{\Pi}}^{-1}(z_1, z_0)$ and the elements of $\tilde{\mathbf{\Pi}}(z_1, z_0)$, the expression for the Fréchet derivatives of \mathbf{M} are derived:

$$\begin{aligned} \mathbf{M}(\Delta\tilde{\mathbf{w}}) = & -(\mathbf{Q}_2^{-1}(z_1, z_0))^T \int_{z_0}^{z_1} \left\{ \mathbf{P}_2^T(z', z_0) \left[\Delta\mathbf{S}_1^{(1)}(\Delta\tilde{\mathbf{w}}) \mathbf{Q}_2(z', z_0) + \Delta\mathbf{S}_1^{(2)}(\Delta\tilde{\mathbf{w}}) \mathbf{P}_2(z', z_0) \right] - \right. \\ & \left. \mathbf{Q}_2^T(z', z_0) \left[\Delta\mathbf{S}_1^{(3)}(\Delta\tilde{\mathbf{w}}) \mathbf{Q}_2(z', z_0) + \Delta\mathbf{S}_1^{(4)}(\Delta\tilde{\mathbf{w}}) \mathbf{P}_2(z', z_0) \right] \right\} dz' \mathbf{Q}_2^{-1}(z_1, z_0) \end{aligned} \quad (\text{B.10})$$

for the derivatives with respect to the initial phase space coordinates and

$$\begin{aligned} \mathbf{M}(\Delta v) = & -(\mathbf{Q}_2^{-1}(z_1, z_0))^T \int_{z_0}^{z_1} \left\{ \mathbf{P}_2^T(z', z_0) \left[\Delta\mathbf{S}_2^{(1)}(\Delta v) \mathbf{Q}_2(z', z_0) + \Delta\mathbf{S}_2^{(2)}(\Delta v) \mathbf{P}_2(z', z_0) \right] - \right. \\ & \left. \mathbf{Q}_2^T(z', z_0) \left[\Delta\mathbf{S}_2^{(3)}(\Delta v) \mathbf{Q}_2(z', z_0) + \Delta\mathbf{S}_2^{(4)}(\Delta v) \mathbf{P}_2(z', z_0) \right] \right\} dz' \mathbf{Q}_2^{-1}(z_1, z_0) \end{aligned} \quad (\text{B.11})$$

for the derivatives with respect to the velocity. The latter integral has to be evaluated for all B-spline coefficients involved in velocity calculations along the central ray under consideration.

Perturbations $\Delta\tilde{\mathbf{\Pi}}(\Delta z_0)$ of the ray propagator matrix due to vertical displacements Δz_0 of the ray starting point are obtained from

$$\begin{aligned} \Delta\tilde{\mathbf{\Pi}}(\Delta z_0) &= \left(\tilde{\mathbf{\Pi}}(z_1, z_0 + \Delta z_0) - \tilde{\mathbf{\Pi}}(z_1, z_0) \right) \\ &= \tilde{\mathbf{\Pi}}(z_1, z_0) \left(\tilde{\mathbf{\Pi}}^{-1}(z_0 + \Delta z_0, z_0) - \tilde{\mathbf{I}} \right) \end{aligned} \quad (\text{B.12})$$

using the first-order Taylor expansion

$$\tilde{\mathbf{\Pi}}(z_0 + \Delta z_0, z_0) \approx \tilde{\mathbf{\Pi}}(z_0, z_0) + \frac{d}{dz} \tilde{\mathbf{\Pi}}(z_0, z_0)|_{z_0} \Delta z_0 = \tilde{\mathbf{I}} + \tilde{\mathbf{S}}|_{z_0} \Delta z_0 \quad (\text{B.13})$$

around z_0 . The matrix $\tilde{\mathbf{S}}$ has been defined in Section 2.4. Inserting the sub-matrices of $\tilde{\mathbf{\Pi}}^{-1}(z_0 + \Delta z_0, z_0)$ into equation (B.12) finally yields

$$\Delta\tilde{\mathbf{\Pi}}(\Delta z_0) = -\tilde{\mathbf{\Pi}}(z_1, z_0) \tilde{\mathbf{S}}|_{z_0} \Delta z_0 . \quad (\text{B.14})$$

The perturbation of matrix \mathbf{M} due to perturbations Δz_0 is consequently expressed as

$$\Delta\mathbf{M} = -(\mathbf{Q}_2^{-1})^T \mathbf{S}^{(2)}|_{z_0} \mathbf{Q}_2^{-1} \Delta z_0 \quad (\text{B.15})$$

using the upper right 2×2 sub-matrix \mathbf{S}^2 of (B.14).

B.1.3 Fréchet derivatives of T

The Fréchet derivatives of traveltime T with respect to the initial phase space coordinates of and the velocity along the central ray are obtained from equations (2.47) and (2.48). Equation (2.48) relates traveltime changes ΔT to changes of the starting position $\underline{\mathbf{x}}_0 = (x_0, y_0, z_0)^T$ of the central ray:

$$\Delta T(\Delta \underline{\mathbf{x}}_0) = -\underline{\mathbf{p}}_0 \cdot \Delta \underline{\mathbf{x}}_0 . \quad (\text{B.16})$$

The Fréchet derivatives with respect to the initial slowness of the central ray vanish due to Fermat's principle. Using equation (A.6) the partial derivatives of traveltime T with respect to the B-spline coefficients are derived from equation (2.47):

$$\frac{\partial T}{\partial v_{jkl}} = \int_{z_0}^{z_1} \frac{1}{v^3 p_z} \beta_j(x) \beta_k(y) \beta_l(-z') dz' . \quad (\text{B.17})$$

The above integral has to be computed along each central ray for all B-spline coefficients involved in the velocity calculations along that ray.

B.2 Fréchet derivatives in ray-centred coordinates

The Fréchet derivatives for 2D NIP-wave tomography and the inversion scheme using model-based diffraction traveltime functions have some transformations at the ray starting and end point in common. They will firstly derived in a notation valid for both inversion schemes. As in the above derivation of the Fréchet derivatives in 3D Cartesian coordinates, quantities at the ray starting point have the index 0 and quantities at the ray end point have the index 1. The Cartesian coordinate axes are named x and z .

For several Fréchet derivatives the relationship between ray-centred and the global Cartesian coordinates is needed. Directly on the central ray, this relation is given by equation (2.61). Let γ denote the angle measured from the ray to the positive vertical axis of the global Cartesian coordinate system. When γ is measured counterclockwise it is defined to be positive. This definition coincides with that used in the description of both inversion schemes. The transformation from ray-centred to Cartesian coordinates reads

$$\begin{pmatrix} q \\ \Delta s \end{pmatrix} = \begin{pmatrix} \cos \gamma & -\sin \gamma \\ \sin \gamma & \cos \gamma \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta z \end{pmatrix} . \quad (\text{B.18})$$

Only infinitesimal small perturbations Δs are allowed in the above transformation. If the magnitude of Δs can not be neglected, local ray-centred Cartesian coordinates need to be introduced at the point under consideration. Taking into account that the measurement surface and, thus, the ray end point is fixed in the inversion schemes, i. e. $\Delta s_1 = \Delta z_1 = 0$, the following relationships are derived from equation (B.18) at the central ray starting and end point:

$$q_0 = \Delta x_0 \cos \gamma_0 - \Delta z_0 \sin \gamma_0 , \quad (\text{B.19a})$$

$$\Delta s_0 = \Delta x_0 \sin \gamma_0 + \Delta z_0 \cos \gamma_0 , \quad (\text{B.19b})$$

$$\Delta x_1 = \frac{1}{\cos \gamma_1} q_1 . \quad (\text{B.19c})$$

In the following, the relationships between slowness changes and perturbations of the take-off and emergence angle of the central ray are derived. For small perturbations $\Delta\gamma$ a perturbed slowness vector $\hat{\mathbf{p}}^{(y)}$ is expressed in local ray-centred Cartesian coordinates (see Section 2.8.1) by

$$\hat{\mathbf{p}}^{(y)} = \frac{1}{v} \begin{pmatrix} \sin(\Delta\gamma) \\ \cos(\Delta\gamma) \end{pmatrix} \approx \frac{1}{v} \begin{pmatrix} \Delta\gamma \\ 1 \end{pmatrix}. \quad (\text{B.20})$$

Taking the difference between this slowness vector and the corresponding unperturbed counterpart yields the perturbation vector

$$\Delta\hat{\mathbf{p}}^{(y)} = \frac{1}{v} \begin{pmatrix} \Delta\gamma \\ 0 \end{pmatrix}. \quad (\text{B.21})$$

The perturbation vector component tangent to the central ray vanishes for small perturbations $\Delta\gamma$. The horizontal component of $\Delta\hat{\mathbf{p}}^{(y)}$ coincides with $p^{(q)}$ used in the paraxial ray tracing system (2.66) in ray-centred coordinates. The above expression is used to establish the relationship between slowness perturbations in ray-centred coordinates due to changes $\Delta\gamma_0$ of the take-off angle of the central ray:

$$p_0^{(q)} = \frac{\Delta\gamma_0}{v_0}. \quad (\text{B.22})$$

At the measurement surface additional considerations are necessary to relate slowness changes in ray-centred coordinates to horizontal slowness changes in global Cartesian coordinates. The unperturbed horizontal component of the slowness vector in the global Cartesian coordinate system is related to the emergence angle γ_1 by $p_{x_1}^{(0)} = v_1^{-1} \sin \gamma_1$. The perturbed horizontal slowness vector component p_{x_1} can be expressed using a first order Taylor expansion around the unperturbed slowness component $p_{x_1}^{(0)}$:

$$p_{x_1} \approx p_{x_1}^{(0)} + \left. \frac{dp_{x_1}^{(0)}}{d\gamma} \right|_{\gamma_1} \Delta\gamma_1 = p_{x_1}^{(0)} + \frac{\cos \gamma_1}{v_1} \Delta\gamma_1. \quad (\text{B.23})$$

Using equation (B.21) and equation (B.23), the final relationship at the measurement surface between slowness perturbations in ray-centred coordinates and perturbations of the horizontal slowness vector component in global Cartesian coordinates is derived:

$$\Delta p_{x_1} = p_1^{(q)} \cos \gamma_1. \quad (\text{B.24})$$

B.2.1 Fréchet derivatives for 2D NIP-wave tomography

The Fréchet derivatives needed in 2D NIP-wave tomography may formally be written as

$$\frac{\partial (T_i, x_{m_i}, p_i, M_i)}{\partial (x_i, z_i, \Theta_i, v_{jk})}, \quad (\text{B.25})$$

with $i = 1, \dots, n_{\text{data}}$, $j = 1, \dots, n_x$, and $k = 1, \dots, n_z$. All derivatives of a specific data point with respect to the NIP-related part of the model description, which corresponds to different data points, vanish. This is not explicitly mentioned in the following derivations where the index indicating the data point is omitted.

Fréchet derivatives of x_m and p

The Fréchet derivatives of x_m and p are derived from equation (2.77) which relates perturbations $\mathbf{w}_1^{(q)}$ of the phase space coordinates at the surface to the corresponding perturbations $\mathbf{w}_0^{(q)}$ at the ray starting point as well as to perturbations of the velocity along the central ray:

$$\mathbf{w}_1^{(q)} = \mathbf{\Pi}^{(q)}(s_1, s_0)\mathbf{w}_0^{(q)} + \mathbf{\Pi}^{(q)}(s_1, s_0) \int_{s_0}^{s_1} \mathbf{\Pi}^{(q)-1}(s', s_0)\Delta\mathbf{n}^{(q)}(s', \Delta v(s')) ds' . \quad (\text{B.26})$$

Using the definition of the reduced Hamiltonian in ray-centred coordinates (2.65), the vector $\Delta\mathbf{n}^{(q)}$, defined in equation (2.78), reads:

$$\Delta\mathbf{n}^{(q)} = \left(0, -\frac{1}{v^2} \frac{\partial \Delta v}{\partial q} + \frac{1}{v^3} \frac{\partial v}{\partial q} \Delta v \right)^T . \quad (\text{B.27})$$

Inserting the normal ray take-off angle Θ , its perturbation $\Delta\Theta$, as well as the normal ray emergence angle β and its associated perturbation $\Delta\beta$ in the expressions derived at the beginning of this section, the Fréchet derivatives of x_m and p with respect to the NIP-related model components x , z , and Θ are as follows:

$$\frac{\partial x_m}{\partial x} = \frac{\cos \Theta}{\cos \beta} Q_1^{(q)}(s_1, s_0) , \quad (\text{B.28a})$$

$$\frac{\partial x_m}{\partial z} = -\frac{\sin \Theta}{\cos \beta} Q_1^{(q)}(s_1, s_0) , \quad (\text{B.28b})$$

$$\frac{\partial x_m}{\partial \Theta} = \frac{1}{v_0 \cos \beta} Q_2^{(q)}(s_1, s_0) , \quad (\text{B.28c})$$

$$\frac{\partial p}{\partial x} = \cos \beta \cos \Theta P_1^{(q)}(s_1, s_0) , \quad (\text{B.28d})$$

$$\frac{\partial p}{\partial z} = -\cos \beta \sin \Theta P_1^{(q)}(s_1, s_0) , \quad (\text{B.28e})$$

$$\frac{\partial p}{\partial \Theta} = \frac{\cos \beta}{v_0} P_2^{(q)}(s_1, s_0) . \quad (\text{B.28f})$$

The Fréchet derivatives of x_m and p with respect to the B-spline coefficients are obtained inserting equation (B.27) into the integral expression in equation (B.26) and considering the two-dimensional analogue of equation (A.6). This finally yields

$$\frac{\partial x_m}{\partial v_{jk}} = \frac{1}{\cos \beta} \left(Q_1^{(q)}(s_1, s_0) \frac{\partial q_1}{\partial v_{jk}} + Q_2^{(q)}(s_1, s_0) \frac{\partial p_1^{(q)}}{\partial v_{jk}} \right) , \quad (\text{B.29a})$$

$$\frac{\partial p}{\partial v_{jk}} = \cos \beta \left(P_1^{(q)}(s_1, s_0) \frac{\partial q_1}{\partial v_{jk}} + P_2^{(q)}(s_1, s_0) \frac{\partial p_1^{(q)}}{\partial v_{jk}} \right) , \quad (\text{B.29b})$$

with

$$\frac{\partial q_1}{\partial v_{jk}} = - \int_{s_0}^{s_1} Q_2^{(q)}(s', s_0) \left\{ -\frac{1}{v^2} \frac{\partial}{\partial q} (\beta_j(x(s'))\beta_k(-z(s'))) + \frac{1}{v^3} \frac{\partial v}{\partial q} (\beta_j(x(s'))\beta_k(-z(s'))) \right\} ds' , \quad (\text{B.30a})$$

$$\frac{\partial p_1^{(q)}}{\partial v_{jk}} = \int_{s_0}^{s_1} Q_1^{(q)}(s', s_0) \left\{ -\frac{1}{v^2} \frac{\partial}{\partial q} (\beta_j(x(s'))\beta_k(-z(s'))) + \frac{1}{v^3} \frac{\partial v}{\partial q} (\beta_j(x(s'))\beta_k(-z(s'))) \right\} ds' . \quad (\text{B.30b})$$

Fréchet derivatives of M

The second traveltime derivative M with respect to the horizontal coordinate at the emergence location of the normal ray is related to the corresponding quantity $M^{(q)}$ in ray-centred coordinates by

$$M = M^{(q)} \cos^2 \beta . \quad (\text{B.31})$$

This transformation requires that the velocity is constant locally around the normal ray emergence location. In the case of strong velocity variations in that region the first partial derivatives of the velocity are involved in the transformation (Červený, 2001). The perturbation of M reads

$$\Delta M = -2 \cos \beta \sin \beta M \Delta \beta + \Delta M^{(q)} \cos^2 \beta \quad (\text{B.32})$$

and depends on perturbations of the normal ray emergence angle β and perturbations of $M^{(q)}$ itself. Writing $M^{(q)}$ in terms of elements of the ray propagator matrix $\mathbf{\Pi}^{(q)}(s_1, s_0)$, the perturbation of $M^{(q)}$ can be expressed in terms of perturbations of the ray propagator matrix elements:

$$\Delta M^{(q)} = \frac{1}{Q_2^{(q)}} \Delta P_2^{(q)} - \frac{P_2^{(q)}}{Q_2^{(q)2}} \Delta Q_2^{(q)} . \quad (\text{B.33})$$

The perturbations of the ray propagator matrix elements are obtained from equation (2.80):

$$\Delta \mathbf{\Pi}^{(q)}(s_1, s_0) = \mathbf{\Pi}^{(q)}(s_1, s_0) \int_{s_0}^{s_1} \mathbf{\Pi}^{(q)-1}(s', s_0) \Delta \mathbf{S}^{(q)}(s') \mathbf{\Pi}^{(q)}(s', s_0) ds' , \quad (\text{B.34})$$

where $\Delta \mathbf{S}^{(q)}$ is divided in two parts depending on perturbations of q and p and on perturbations of the velocity, respectively. For the same reason outlined in the derivation of the 3D Fréchet derivatives in Cartesian coordinates, only the upper right element of the integral expression in equation (B.34) needs actually to be evaluated. This leads to the following expressions for perturbations $\Delta M^{(q)}$ due to perturbations of the ray propagator matrix elements:

$$\begin{aligned} \Delta M^{(q)}(\Delta v, \Delta q_0, \Delta p_0^{(q)}) = & -Q_2^{(q)-2}(s_1, s_0) \int_{s_0}^{s_1} \left\{ P_2^{(q)}(s', s_0) Q_2^{(q)}(s', s_0) \left[\Delta S_{11}^{(q)}(s') - \Delta S_{22}^{(q)}(s') \right] \right. \\ & \left. + P_2^{(q)2}(s', s_0) \Delta S_{12}^{(q)} - Q_2^{(q)2}(s', s_0) \Delta S_{21}^{(q)}(s') \right\} ds' . \end{aligned} \quad (\text{B.35})$$

The perturbation of the ray propagator matrix due to a perturbation Δs_0 is obtained in a way analogous to that used to derive the perturbation of the ray propagator matrix in Cartesian coordinates due to a vertical displacement of the ray starting point. For the calculation of $\Delta M^{(q)}(\Delta s_0)$, only the upper right element needs to be evaluated. The expression finally reads

$$\Delta M^{(q)}(\Delta s_0) = Q_2^{(q)-2}(s_1, s_0) v(s_0) \Delta s_0 . \quad (\text{B.36})$$

The final expressions for the Fréchet derivatives of M are derived using the relationships outlined at the beginning of this section and equations (B.28):

$$\begin{aligned} \frac{\partial M}{\partial x} = & -2 \sin \beta \cos \beta \cos \Theta v_1 P_1^{(q)}(s_1, s_0) \frac{P_2^{(q)}(s_1, s_0)}{Q_2^{(q)}(s_1, s_0)} \\ & + \cos^2 \beta \left(\frac{\partial M^{(q)}}{\partial q_0} \cos \Theta + Q_2^{(q)-2}(s_1, s_0) v_0 \sin \Theta \right), \end{aligned} \quad (\text{B.37a})$$

$$\begin{aligned} \frac{\partial M}{\partial z} = & 2 \sin \beta \cos \beta \sin \Theta v_1 P_1^{(q)}(s_1, s_0) \frac{P_2^{(q)}(s_1, s_0)}{Q_2^{(q)}(s_1, s_0)} \\ & + \cos^2 \beta \left(-\frac{\partial M^{(q)}}{\partial q_0} \sin \Theta + Q_2^{(q)-2}(s_1, s_0) v_0 \cos \Theta \right), \end{aligned} \quad (\text{B.37b})$$

$$\frac{\partial M}{\partial \Theta} = -2 \sin \beta \cos \beta \frac{v_1 P_2^{(q)2}(s_1, s_0)}{v_0 Q_2^{(q)}(s_1, s_0)} + \frac{1}{v_0} \frac{\partial M^{(q)}}{\partial p_0}, \quad (\text{B.37c})$$

$$\frac{\partial M}{\partial v_{jk}} = -2 \sin \beta \cos \beta \frac{P_2^{(q)}(s_1, s_0)}{Q_2^{(q)}(s_1, s_0)} \left(P_1^{(q)}(s_1, s_0) \frac{\partial q_1}{\partial v_{jk}} + P_2^{(q)}(s_1, s_0) \frac{\partial p_1}{\partial v_{jk}} \right) + \frac{\partial M^{(q)}}{\partial v_{jk}}. \quad (\text{B.37d})$$

The expressions for $\partial M^{(q)}/\partial q_0$, $\partial M^{(q)}/\partial p_0$, and $\partial M^{(q)}/\partial v_{jk}$ are derived from equation (B.35) inserting the corresponding expression for $\Delta S^{(q)}$.

Fréchet derivatives of T

The Fréchet derivatives of T are obtained from equation (2.79) which reads in the here used notation

$$\Delta T = - \int_{s_0}^{s_1} \frac{\Delta v}{v^2(s')} ds' - \frac{\sin \Theta}{v_0} \Delta x_0 - \frac{\cos \Theta}{v_0} \Delta z_0. \quad (\text{B.38})$$

Considering the two-dimensional analogue of equation (A.6) they are given by

$$\frac{\partial T}{\partial x} = -\frac{\sin \Theta}{v_0}, \quad (\text{B.39a})$$

$$\frac{\partial T}{\partial z} = -\frac{\cos \Theta}{v_0}, \quad (\text{B.39b})$$

$$\frac{\partial T}{\partial \Theta} = 0, \quad (\text{B.39c})$$

$$\frac{\partial T}{\partial v_{jk}} = - \int_{s_0}^{s_1} \frac{\beta_j(x(s')) \beta_k(-z(s'))}{v^2(s')} ds' \quad (\text{B.39d})$$

The partial derivatives of T with respect to Θ for fixed starting and end points vanishes due to Fermat's principle. For the same reason, perturbations of the starting and end points perpendicular to the ray have no first-order effect on the traveltime along the central ray trajectory.

B.2.2 Fréchet derivatives for the inversion scheme using model-based diffraction traveltime functions

In the inversion scheme using model-based diffraction traveltime functions as moveout operators, the needed Fréchet derivatives can formally be expressed by (see section 4.3.4)

$$\frac{\partial \left(T_{0_i}, p_{0_i}, x_{m_{0_i}}, t_1(\Theta_{0_i}, \Phi_l), \frac{\partial}{\partial \Theta} t_i(\Theta_{0_i}, \Phi_l) \right)}{\partial (x_i, z_i, \Theta_{0_i}, v_{jk})}, \quad (\text{B.40})$$

with $i = 1, \dots, n_{\text{data}}$, $l = 1, \dots, n_{\text{pick}}$, $j = 1, \dots, n_x$, and $k = 1, \dots, n_z$. The ZO data components have an index 0 in the above expression. Their Fréchet derivatives with respect to the model parameters have already been outlined in this section and are not repeated here. The normal ray take-off angle, which coincides with the geological dip estimate at the reflection point, is denoted by Θ_0 in equation (B.40) to distinguish it from all hypothetical geological dips involved in the modelling of the diffraction traveltime function. It coincides with Θ used in the above derivation of the Fréchet derivatives for 2D NIP-wave tomography. In the following, all subscripts indicating the pick or data point number are omitted. All Fréchet derivatives

$$\frac{\partial \left(T_{0_i}, p_{0_i}, x_{m_{0_i}}, t_i(\Theta_{0_i}, \Phi_l), \frac{\partial}{\partial \Theta} t_i(\Theta_{0_i}, \Phi_l) \right)}{\partial (x_j, z_j, \Theta_{0_j})}, \quad i \neq j,$$

vanish.

The partial derivatives of $t(\Theta_0, \Phi)$ are obtained considering the decomposition of the total traveltime into a source and a receiver part, i. e., $t = t_S + t_G$. The dependence on (x, z, Θ_0) of the traveltime along the two rays which connect the source and the receiver with the reflection point is described by expressions analogous to those derived above for normal rays. The Fréchet derivatives of $t(\Theta_0, \Phi)$ are given by the sum of the individual terms related to the two ray segments:

$$\frac{\partial t(\Theta_0, \Phi)}{\partial x} = -\frac{\sin \gamma_S}{v_0} - \frac{\sin \gamma_G}{v_0}, \quad (\text{B.41a})$$

$$\frac{\partial t(\Theta_0, \Phi)}{\partial z} = -\frac{\cos \gamma_S}{v_0} - \frac{\cos \gamma_G}{v_0}, \quad (\text{B.41b})$$

$$\frac{\partial t(\Theta_0, \Phi)}{\partial \Theta_0} = 0, \quad (\text{B.41c})$$

$$\frac{\partial t(\Theta_0, \Phi)}{\partial v_{jk}} = -\int_{s_{S_0}}^{s_{S_1}} \frac{\beta_j(x(s'_S))\beta_k(-z(s'_S))}{v^2(s'_S)} ds'_S - \int_{s_{G_0}}^{s_{G_1}} \frac{\beta_j(x(s'_G))\beta_k(-z(s'_G))}{v^2(s'_G)} ds'_G, \quad (\text{B.41d})$$

where γ_S , γ_G , s_S , and s_G denote the take-off angles of and arclength along the source and receiver rays, respectively.

The angle domain time dip $\frac{\partial}{\partial \Theta} t(\Theta_0, \Phi)$ is related to the horizontal slowness at the source and at the receiver according to equation (4.12):

$$\frac{\partial t}{\partial \Theta} = p_S \frac{\partial x_S}{\partial \gamma_S} + p_G \frac{\partial x_G}{\partial \gamma_G}. \quad (\text{B.42})$$

To simplify the notation in the following expressions, the angle domain time dip $\frac{\partial}{\partial \Theta} t(\Theta_0, \Phi)$ is denoted by $t_{,\Theta}$. The model dependent transformation factors $\partial x_S / \partial \gamma_S$ and $\partial x_G / \partial \gamma_G$ between surface

coordinates and take-off angles are assumed to be constant in the following derivations. Their dependence on the reflection point position, the geological dip, as well as on the velocity distribution along the source and receiver rays is neglected in the calculation of the Fréchet derivatives. Including these dependencies requires to consider perturbations of the propagator matrices along both rays. In the tomographic inversion, the transformation factors change from iteration to iteration as the model changes. Using equation (B.28) the relationship between the angle domain time dip and the horizontal slowness at the source and at the receiver reads

$$t_{,\Theta} = p_S \frac{Q_2^{(q)}(s_{S_1}, s_{S_0})}{v_0 \cos \beta_S} + p_G \frac{Q_2^{(q)}(s_{G_1}, s_{G_0})}{v_0 \cos \beta_G}, \quad (\text{B.43})$$

where β_S and β_G denote the emergence angle of the source and receiver rays at the surface and $Q_2^{(q)}(s_{S_1}, s_{S_0})$ and $Q_2^{(q)}(s_{G_1}, s_{G_0})$ are elements of the propagator matrices along these rays. Using the expressions (B.28) and equation (4.1), the Fréchet derivatives of $t_{,\Theta}$ are given by

$$\frac{\partial}{\partial x} t_{,\Theta} = \frac{\cos \gamma_S}{v_0} P_1^{(q)}(s_{S_1}, s_{S_0}) Q_2^{(q)}(s_{S_1}, s_{S_0}) + \frac{\cos \gamma_G}{v_0} P_1^{(q)}(s_{G_1}, s_{G_0}) Q_2^{(q)}(s_{G_1}, s_{G_0}), \quad (\text{B.44a})$$

$$\frac{\partial}{\partial z} t_{,\Theta} = -\frac{\sin \gamma_S}{v_0} P_1^{(q)}(s_{S_1}, s_{S_0}) Q_2^{(q)}(s_{S_1}, s_{S_0}) - \frac{\sin \gamma_G}{v_0} P_1^{(q)}(s_{G_1}, s_{G_0}) Q_2^{(q)}(s_{G_1}, s_{G_0}), \quad (\text{B.44b})$$

$$\frac{\partial}{\partial \Theta_0} t_{,\Theta} = \frac{1}{v_0^2} P_2^{(q)}(s_{S_1}, s_{S_0}) Q_2^{(q)}(s_{S_1}, s_{S_0}) + \frac{1}{v_0^2} P_2^{(q)}(s_{G_1}, s_{G_0}) Q_2^{(q)}(s_{G_1}, s_{G_0}), \quad (\text{B.44c})$$

$$\begin{aligned} \frac{\partial}{\partial v_{jk}} t_{,\Theta} = & \left(P_1^{(q)}(s_{S_1}, s_{S_0}) \frac{\partial q_{S_1}}{\partial v_{jk}} + P_2^{(q)}(s_{S_1}, s_{S_0}) \frac{\partial p_{S_1}^{(q)}}{\partial v_{jk}} \right) \frac{Q_2^{(q)}(s_{S_1}, s_{S_0})}{v_0} \\ & + \left(P_1^{(q)}(s_{G_1}, s_{G_0}) \frac{\partial q_{G_1}}{\partial v_{jk}} + P_2^{(q)}(s_{G_1}, s_{G_0}) \frac{\partial p_{G_1}^{(q)}}{\partial v_{jk}} \right) \frac{Q_2^{(q)}(s_{G_1}, s_{G_0})}{v_0}. \end{aligned} \quad (\text{B.44d})$$

B.3 Fréchet derivatives for additional constraints

The additional constraints that can be applied to stabilise the inversion schemes have been discussed in Chapter 5. In the inversion schemes they are handled as additional components in the data vector. The associated Fréchet derivatives are derived in this section.

B.3.1 A priori velocity information

The forward-modelled counterpart to a priori known velocity values at locations $\mathbf{x}_i = (x_i, z_i)^T$, $i = 1, \dots, n_{\text{data}}$, is obtained evaluating the B-spline velocity model parameterisation at these locations. For the general 3D case, this reads

$$v(x_i, y_i, z_i) = \sum_{j=1}^{n_x} \sum_{k=1}^{n_y} \sum_{l=1}^{n_z} v_{jkl} \beta_j(x_i) \beta_k(y_i) \beta_l(-z_i), \quad (\text{B.45})$$

with $j = 1, \dots, n_x$, $k = 1, \dots, n_y$, and $l = 1, \dots, n_z$. Obviously, the above expression only depends on the B-spline coefficients. Thus, the Fréchet derivatives with respect to the NIP-related part of the

model description vanish. The partial derivatives of the above expression with respect to the B-spline coefficients v_{jkl} read

$$\frac{\partial v(x_i, y_i, z_i)}{\partial v_{jkl}} = \beta_j(x_i) \beta_k(y_i) \beta_l(-z_i). \quad (\text{B.46})$$

The corresponding expressions in the 2D case are obtained omitting the y -component in equation (B.46).

B.3.2 Minimum velocity variation along reflectors

As described in Chapter 5 the velocity structure can locally be forced to follow the reflector structure by minimising the component of the gradient vertical to the normal rays at the NIPs. This constraint depends on the velocity model and on the NIP-related part of the model description. However, for an efficient implementation the dependence on the NIP-related part of the model description is assumed to be small and is neglected. This, of course, only holds when the change in the NIP-locations and the associated normal ray take-off angles is small from iteration to iteration. This can be assured using a small update increment λ (see Chapter 5). The Fréchet derivatives of this constraint have to be evaluated for all NIP-locations, i. e. for all data points. In the following derivation the index i for the data point is omitted. All calculations are directly performed at the NIP locations.

As already outlined in Chapter 5, the projection of the local velocity gradient into the plane orthogonal to the normal ray reads

$$\nabla_{\perp} v = \left(\hat{\mathbf{e}}_1 \cdot \underline{\mathbf{e}}^{(v)}, \hat{\mathbf{e}}_2 \cdot \underline{\mathbf{e}}^{(v)} \right)^T, \quad (\text{B.47})$$

with

$$\hat{\mathbf{e}}_1 \cdot \underline{\mathbf{e}}^{(v)} = \sum_{j=1}^{n_x} \sum_{k=1}^{n_y} \sum_{l=1}^{n_z} v_{jkl} \left(\hat{e}_{1,x} \frac{\partial \beta_j(x)}{\partial x} \beta_k(y) \beta_l(-z) + \hat{e}_{1,y} \beta_j(x) \frac{\partial \beta_k(y)}{\partial y} \beta_l(-z) - \hat{e}_{1,z} \beta_j(x) \beta_k(y) \frac{\partial \beta_l(-z)}{\partial z} \right) \quad (\text{B.48})$$

and an analogous expression for $\hat{\mathbf{e}}_2 \cdot \underline{\mathbf{e}}^{(v)}$. The magnitude of the velocity gradient projection is given by

$$|\nabla_{\perp} v| = \sqrt{(\hat{\mathbf{e}}_1 \cdot \underline{\mathbf{e}}^{(v)})^2 + (\hat{\mathbf{e}}_2 \cdot \underline{\mathbf{e}}^{(v)})^2}. \quad (\text{B.49})$$

The first partial derivatives of equation (B.49) with respect to the B-spline coefficients v_{jkl} are derived using equation (B.48) and its analogue for $\hat{\mathbf{e}}_2 \cdot \underline{\mathbf{e}}^{(v)}$:

$$\frac{\partial}{\partial v_{jkl}} |\nabla_{\perp} v| = \frac{1}{|\nabla_{\perp} v|} \left((\hat{\mathbf{e}}_1 \cdot \underline{\mathbf{e}}^{(v)}) \frac{\partial}{\partial v_{jkl}} (\hat{\mathbf{e}}_1 \cdot \underline{\mathbf{e}}^{(v)}) + (\hat{\mathbf{e}}_2 \cdot \underline{\mathbf{e}}^{(v)}) \frac{\partial}{\partial v_{jkl}} (\hat{\mathbf{e}}_2 \cdot \underline{\mathbf{e}}^{(v)}) \right), \quad (\text{B.50})$$

with

$$\frac{\partial}{\partial v_{jkl}} (\hat{\mathbf{e}}_1 \cdot \underline{\mathbf{e}}^{(v)}) = \left(\hat{e}_{1,x} \frac{\partial \beta_j(x)}{\partial x} \beta_k(y) \beta_l(-z) + \hat{e}_{1,y} \beta_j(x) \frac{\partial \beta_k(y)}{\partial y} \beta_l(-z) - \hat{e}_{1,z} \beta_j(x) \beta_k(y) \frac{\partial \beta_l(-z)}{\partial z} \right) \quad (\text{B.51})$$

and an expression of the same form for $\frac{\partial}{\partial v_{jkl}} (\hat{\mathbf{e}}_2 \cdot \underline{\mathbf{e}}^{(v)})$. The corresponding Fréchet derivatives for the 2D case are derived from the above expressions choosing $\hat{\mathbf{e}}_1$ in the (x, z) -plane and omitting all terms related to the y -component.

List of Figures

1.1	Definition of midpoint and half-offset coordinates.	3
2.1	Ray-centred coordinates of a point in the vicinity of a central ray.	21
3.1	Hypothetical NIP- and normal-wavefronts.	30
3.2	Definition of the event-aligned volume.	32
3.3	Stacking velocity calculated from original and smoothed kinematic wavefield attributes along an identified reflection event in a CRS stacked ZO volume.	33
3.4	Smoothed values and uncertainty estimates for the horizontal slowness and the second traveltme derivative using the robust linear regression technique.	34
3.5	The velocity model building work flow based on second-order approximations of diffraction traveltme functions.	45
3.6	A selected CMP gather of the synthetic seismic prestack data set with and without random noise	47
3.7	CRS stack section and associated coherence after simultaneous three-parameter optimisation.	48
3.8	CRS stacking parameters M and p before and after application of the event-consistent smoothing using the local robust linear regression technique.	49
3.9	The final smooth velocity model obtained after 12 iterations and the original blocky velocity model.	50
3.10	Stack of prestack depth migrated CO sections and selected CIGs.	51
4.1	Relationship between time and depth domain for the ZO case.	55
4.2	Definition of the illumination angle Θ and the scattering angle Φ	56
4.3	Traveltme distributions in the correct velocity model for an analytical example after the moveout correction with the model-based diffraction traveltme operator.	57
4.4	Time dip distributions in the correct velocity model for an analytical example after the moveout correction with the model-based diffraction traveltme operator.	58

4.5	Traveltime distributions in a wrong velocity model for an analytical example after the moveout correction with the model-based diffraction traveltime operator.	59
4.6	Time dip distributions in a wrong velocity model for an analytical example after the moveout correction with the model-based diffraction traveltime operator.	60
4.7	The CRP gather and a common-scattering-angle gather extracted from the moveout corrected prestack data cube along with the definition of traveltime residuals $\Delta t(\Phi)$ and time dip residuals $\frac{\partial}{\partial \Phi} \Delta t(\Phi)$	63
4.8	Single pick inversion: results are obtained using only traveltime residuals picked in a CIG.	70
4.9	Single pick inversion: results are obtained using only traveltime residuals picked in a CRP gather.	71
4.10	Single pick inversion: results are obtained using traveltime residuals picked in a CRP gather and time dip residuals.	72
4.11	ZO stack section with 412 selected pick locations.	73
4.12	Inverted reflection point locations with the true, blocky velocity model in the background.	74
4.13	Inverted smooth velocity model overlaid with the stack of prestack depth migrated common-offset sections.	74
4.14	The stack of prestack depth migrated common-offset sections using the inverted velocity model with the true, blocky velocity model in the background.	75
4.15	Selected CIGs. A maximum offset of 4000 m is displayed.	75
6.1	Smooth velocity model used for modelling of the seismic prestack data and the depth-migrated forward-modelled ZO section.	84
6.2	A selected CMP gather.	84
6.3	CRS stack results.	86
6.4	Automatically detected pick locations on top of the CRS stack section.	86
6.5	Velocity models obtained by NIP-wave tomography and the update based on forward-modelled diffraction traveltime functions.	87
6.6	Stack of prestack depth migrated common-offset sections. Migration has been performed using the final model of NIP-wave tomography and the model after the update with model-based diffraction traveltime functions.	88
6.7	Selected CIG obtained with the final velocity model of NIP-wave tomography and the updated velocity model.	89
6.8	Azimuthal distribution of offsets.	90
6.9	ZO trace locations. Highlighted in red are two cross sections for which the CRS results are shown.	91

6.10	CRS stack and coherence sections for profile 1 and profile 2.	91
6.11	CRS attribute sections (normal ray emergence angle and azimuth) for profile 1 and profile 2.	92
6.12	CRS attribute sections (elements of matrix \mathbf{M}) for profile 1 and profile 2.	93
6.13	Difference in stacking velocity calculated for the x-direction and y-direction along profile 1 and 2.	94
6.14	Pick locations.	95
6.15	Picks depicted as stacking velocity versus ZO one-way time for three azimuthal directions.	96
6.16	Slices through the velocity models obtained with the single azimuth and the multi-azimuth inversion.	97
6.17	Final residual errors of the one-way traveltimes t and the horizontal slowness vector components as a function of pick location for the single-azimuth and the multi-azimuth inversion.	98
6.18	Final residual errors for the elements of matrix \mathbf{M} as a function of pick location for the single-azimuth and multi-azimuth (right) inversion.	99
6.19	Lateral distance between the NIP-locations obtained with single-azimuth and multi-azimuth inversion as a function of pick location.	100
6.20	Definition of offset azimuth bins.	100
6.21	Azimuthally dependent CIGs at $x = 6500$ m obtained with the velocity models from the single-azimuth and the multi-azimuth inversion.	102
6.22	Azimuthally dependent CIGs at $y = 2500$ m obtained with the velocity models from a) the single-azimuth and b) the multi-azimuth inversion.	103
6.23	Depth-migrated CRS stack sections using the velocity model from the single-azimuth and the multi-azimuth inversions.	104
B.1	Relationship between horizontal and vertical displacements at the starting position of the central ray.	112

References

- Aki, K. and Richards, P. G. (1980). *Quantitative seismology – theory and methods*, volume 1. W. H. Freeman & Co., San Francisco.
- Al-Chalabi, M. (1973). Series approximation in velocity and travelttime computations. *Geophys. Prosp.*, 21:783–795.
- Al-Yahya, K. (1989). Velocity analysis by iterative profile migration. *Geophysics*, 54(6):718–729.
- Audebert, F. and Diet, J. P. (1993). Migrated focus panels: Focusing analysis reconciled with prestack depth migration. In *Expanded abstracts, 63rd Ann. Internat. Mtg.*, pages 961–964. Soc. Expl. Geophys.
- Billette, F. and Lambaré, G. (1998). Velocity macro-model estimation from seismic reflection data by stereotomography. *Geophys. J. Intern.*, 135:671–690.
- Billette, F., Le Bégat, S., Podvin, P., and Lambaré, G. (2003). Practical aspects and applications of 2D stereotomography. *Geophysics*, 68(3):1008–1021.
- Bishop, T., Bube, K., Cutler, R., Langan, R., Love, P., Resnick, J., Shuey, R., Spindler, D., and Wyld, H. (1985). Tomographic determination of velocity and depth in laterally varying media. *Geophysics*, 50(1):903–923.
- Bleistein, N. (1984). *Mathematical methods for wave phenomena*. Academic Press Inc., San Diego.
- Červený, V. (2001). *Seismic ray theory*. Cambridge University Press, Cambridge.
- Chauris, H., Noble, M., Lambaré, G., and Podvin, P. (2002a). Migration velocity analysis from locally coherent events in 2-D laterally heterogeneous media, part I: theoretical aspects. *Geophysics*, 67(4):1202–1212.
- Chauris, H., Noble, M., Lambaré, G., and Podvin, P. (2002b). Migration velocity analysis from locally coherent events in 2-D laterally heterogeneous media, part II: applications on synthetic and real data. *Geophysics*, 67(4):1213–1224.
- Chernyak, V. S. and Gritsenko, S. A. (1979). Interpretation of the effective common-depth-point parameters for a three-dimensional system of homogeneous layers with curvilinear boundaries. *Geologiya i Geofizika*, 20(12):112–120.
- de Boor, C. (1978). *A practical guide to splines*. Springer-Verlag, Berlin.

References

- Della Moretta, D., Klüver, T., and Marchetti, P. (2006). 3D CRS-Based Velocity Model Building – An Accurate and Cost-Effective Approach. In *Extended abstracts, 68th Conf. Eur. Assn. Geosci. Eng.*
- Delprat-Jannaud, F. and Lailly, P. (1993). Ill-posed and well-posed formulations of the reflection travel time tomography problem. *J. Geophys. Res.*, 98(B4):6589–6605.
- Deregowski, S. M. (1990). Common-offset migrations and velocity analysis. *First Break*, 8(6):224–234.
- Dix, C. H. (1955). Seismic velocities from surface measurements. *Geophysics*, 20(1):68–86.
- Duveneck, E. (2003). Determination of velocity models from data-derived wavefront attributes. In *Extended abstracts, 65th Conf. Eur. Assn. Geosci. Eng. Session D27.*
- Duveneck, E. (2004a). *Tomographic determination of seismic velocity models with kinematic wavefield attributes.* Logos Verlag, Berlin.
- Duveneck, E. (2004b). Velocity model estimation with data-derived wavefront attributes. *Geophysics*, 69(1):265–274.
- Duveneck, E. and Hubral, P. (2002). Tomographic velocity model inversion using kinematic wavefield attributes. In *Expanded abstracts, 72nd Ann. Internat. Mtg.*, pages 862–865. Soc. Expl. Geophys.
- Farra, V. and Madariaga, R. (1987). Seismic waveform modeling in heterogeneous media by ray perturbation theory. *J. Geophys. Res.*, 92(B3):2697–2712.
- Farra, V. and Madariaga, R. (1988). Non-linear reflection tomography. *Geophys. J.*, 95:135–147.
- Gilbert, F. and Backus, G. (1966). Propagator matrices in elastic wave and vibration problems. *Geophysics*, 31(2):326–332.
- Gjøystdal, H. and Ursin, B. (1981). Inversion of reflection times in three dimensions. *Geophysics*, 46(7):972–983.
- Hertweck, T., Jäger, C., Mann, J., and Duveneck, E. (2003). An integrated data-driven approach to seismic reflection imaging. In *Extended abstracts, 65th Conf. Eur. Assn. Geosci. Eng. Session P004.*
- Hertweck, T., Mann, J., and Klüver, T. (2005). Event-consistent smoothing in the context of the crs stack method. *J. Seis. Expl.*, 14:197–215.
- Höcht, G. (2002). *Traveltime approximations for 2D and 3D media and kinematic wavefield attributes.* PhD thesis, University of Karlsruhe.
- Hubral, P. (1983). Computing true amplitude reflections in a laterally inhomogeneous earth. *Geophysics*, 48(8):1051–1062.
- Hubral, P. and Krey, T. (1980). *Interval velocities from seismic reflection traveltime measurements.* Soc. Expl. Geophys., Tulsa.
- Jäger, R., Mann, J., Höcht, G., and Hubral, P. (2001). Common-Reflection-Surface stack: image and attributes. *Geophysics*, 66(1):97–109.

- Jeannot, J. P. and Faye, J. P. (1986). Prestack migration velocities from focusing depth analysis. In *Expanded abstracts, 56th Ann. Internat. Mtg.*, pages 438–440. Soc. Expl. Geophys.
- Kravtsov, Y. A. and Orlov, Y. I. (1990). *Geometrical optics of inhomogeneous media*. Springer Verlag, Berlin.
- Lafond, C. F. and Levander, A. R. (1993). Migration moveout analysis and depth focusing. *Geophysics*, 58(1):91–100.
- Lavaud, B., Baina, R., and Landa, E. (2004). Poststack Stereotomography – A Robust Strategy for Velocity Model Estimation. In *Extended abstracts, 66th Conf. Eur. Assn. Geosci. Eng. Session C022*.
- Lines, L. R. and Treitel, S. (1984). Tutorial: a review of least-squares inversion and its application to geophysical problems. *Geophys. Prosp.*, 32:159–186.
- Liu, Z. and Bleistein, N. (1995). Migration velocity analysis: theory and an iterative algorithm. *Geophysics*, 60(1):142–153.
- MacKay, S. and Abma, R. (1992). Imaging and velocity estimation with depth-focusing analysis. *Geophysics*, 57(12):1608–1622.
- Mann, J. (2002). *Extensions and applications of the Common-Reflection-Surface Stack method*. Logos Verlag, Berlin.
- Mann, J. and Höcht, G. (2003). Pulse stretch effects in the context of data-driven imaging methods. In *Extended abstracts, 65th Conf. Eur. Assn. Geosci. Eng. Session P007*.
- Mann, J., Jäger, R., Müller, T., Höcht, G., and Hubral, P. (1999). Common-Reflection-Surface stack – a real data example. *J. Appl. Geoph.*, 42(3,4):301–318.
- Mayne, W. H. (1962). Common reflection point horizontal data stacking techniques. *Geophysics*, 27(6):927–938.
- Menke, W. (1984). *Geophysical data analysis: discrete inverse theory*. Academic Press, Orlando.
- Müller, A. (2003). The 3D Common-Reflection-Surface Stack - Theory and Application. Master's thesis, University of Karlsruhe.
- Müller, A. (2006). Elimination of the spread-length bias in the Common-Reflection-Surface stack. In *Expanded abstracts, 76th Ann. Internat. Mtg.*, pages 3006–3010. Soc. Expl. Geophys.
- Müller, T. (1999). *The Common Reflection Surface stack method – seismic imaging without explicit knowledge of the velocity model*. Der Andere Verlag, Bad Iburg.
- Neckludov, D., Baina, R., and Landa, E. (2006). Residual stereotomographic inversion. *Geophysics*, 71(4):E35–E39.
- Neidell, N. S. and Taner, M. T. (1971). Semblance and other coherency measures for multichannel data. *Geophysics*, 36(3):482–497.
- Paige, C. C. and Saunders, M. A. (1982a). Algorithm 583 – LSQR: sparse linear equations and least squares problems. *ACM Trans. Math. Softw.*, 8(2):195–209.

References

- Paige, C. C. and Saunders, M. A. (1982b). LSQR: an algorithm for sparse linear equations and sparse least squares. *ACM Trans. Math. Softw.*, 8(1):43–71.
- Popov, M. M. and Psencik, I. (1978). Computation of ray amplitudes in inhomogeneous media with curved interfaces. *Studia Geoph. et Geod.*, 22:248–258.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (1992). *Numerical recipes in C; the art of scientific computing*. Cambridge University Press, Cambridge, 2nd edition.
- Robein, E. (2003). *Velocities, Time-imaging and Depth-imaging in Reflection Seismics. Principles and Methods*. EAGE Publications, Houten.
- Sattlegger, J. W., Rohde, J., Egbers, H., Dohr, G. P., Stiller, P. K., and Echterhoff, J. A. (1981). INMOD – two dimensional inverse modeling algorithm based on ray theory. *Geophys. Prosp.*, 29:229–240.
- Schneider, W. (1978). Integral formulation for migration in two and three dimensions. *Geophysics*, 43(1):49–76.
- Söllner, W. and Yang, W.-Y. (2002). Diffraction response simulation: A tool for velocity determination. In *Expanded abstracts, 66th Ann. Internat. Mtg.*, pages 1172–1175. Soc. Expl. Geophys.
- Stork, C. (1992). Reflection tomography in the postmigrated domain. *Geophysics*, 57(5):680–692.
- Stork, C. and Clayton, R. W. (1991). Linear aspects of tomographic velocity analysis. *Geophysics*, 56(4):483–495.
- Taner, M. T. and Koehler, F. (1969). Velocity spectra – digital computer derivation and applications of velocity functions. *Geophysics*, 34(6):859–881.
- Tarantola, A. (1987). *Inverse problem theory: methods for data fitting and model parameter estimation*. Elsevier, Amsterdam.
- Trappe, H., Mann, J., and Buske, S. (2005). High-resolution images of subsurface CO₂ storage sites in time and depth by the CRS methodology (CO₂CRS). In *Geotechnologien Science Report No. 6: Investigation, utilization and protection of the underground*. Koordinierungsbüro Geotechnologien, Ext. Abstr., Potsdam.
- Williamson, P. (1990). Tomographic inversion in reflection seismology. *Geophys. J. Intern.*, 100:255–274.
- Woodward, M., Farmer, P., Nichols, D., and Charles, S. (1998). Automated 3D tomographic velocity analysis of residual moveout in prestack migrated common image point gathers. In *Expanded abstracts, 68th Ann. Internat. Mtg.*, pages 1218–1221. Soc. Expl. Geophys.
- Yilmaz, Ö. (1987). *Seismic data processing*. Soc. Expl. Geophys., Tulsa.
- Yilmaz, Ö. and Chambers, R. (1984). Migration velocity analysis by wavefield extrapolation. *Geophysics*, 49(10):1664–1674.

Danksagung

Mein besonderer Dank gilt **Prof. Dr. Peter Hubral** für die Betreuung meiner Arbeit. Er hat mich in jeder Hinsicht unterstützt und es mir ermöglicht, meine Arbeit im Rahmen des WIT-Konsortiums sowie auf internationalen Konferenzen vorzustellen.

Prof. Dr. Dirk Gajewski danke ich für die Übernahme des Korreferates und sein Interesse an meiner Arbeit.

Ich möchte mich außerdem bei allen ehemaligen und jetzigen Kollegen bedanken, mit denen ich in der Arbeitsgruppe Reflexionsseismik zusammengearbeitet habe.

Besonderer Dank gilt **Dr. Eric Duveneck**, auf dessen Arbeit ich hier aufbaue. Er hat mich während meiner Diplomarbeit sehr gut betreut und mir den Zugang zur reflexionsseismischen Tomographie ermöglicht.

Großer Dank gebührt **Dr. Jürgen Mann, Alex Müller** und **Miriam Spinner** für gute Zusammenarbeit, viele hilfreiche Diskussionen sowie das Korrekturlesen dieser Arbeit.

Den Sponsoren des Wave-Inversion-Technology (WIT) Konsortiums danke ich für ihre Unterstützung.

Dr. Henning Trappe und **Dr. Jürgen Pruessmann** von *TEEC* danke ich für ihr anhaltendes Interesse an der Tomographie.

Thanks to **Dr. Thomas Hertweck** (*Fugro*) and his colleagues for the excellent cooperation.

I am grateful to **Paolo Marchetti** and **Davide della Moretta** (*ENI E&P*) for the good cooperation.

Thanks a lot to **Dr. Evgeny Landa** for his invitations to visit the *OPERA* group in Pau. Thank you also to all other members of the *OPERA* group, especially to **Dr. Reda Baina** and **German Höcht**, for good cooperation and a great time. My visits in Pau and the discussions I had there contributed a lot to this work. Thank you also for providing the synthetic 2D dataset I used in Chapter 6.

Allen Mitarbeitern des Geophysikalischen Instituts danke ich für eine angenehme Zeit. Besonderer Dank gebührt **Claudia Payne** für ihre Hilfe bei verschiedenen Gelegenheiten.

Mein besonderer Dank gilt **Cathérine Meißner**, nicht nur für das Korrekturlesen dieser Arbeit.

Meinen Eltern danke ich für die Unterstützung nicht nur während des Studiums, die die Entstehung dieser Arbeit erst möglich gemacht hat.

Lebenslauf

Persönliche Daten

Name:	Tilman Klüver
Geburtsdatum:	17. September 1977
Nationalität:	deutsch
Geburtsort:	Lüneburg

Schulbildung

1984-1988	Grundschule Lüne, Lüneburg
1988-1990	Orientierungsstufe Lüne, Lüneburg
1990-1997	Gymnasium Wilhelm-Raabe-Schule, Lüneburg
10.06.1997	Abitur

Wehrdienst

1997-1998	Grundwehrdienst
-----------	-----------------

Hochschulbildung

1998 - 2004	Studium der Geophysik an der Universität Karlsruhe (TH)
01.03.2004	Diplom
seit 2004	Doktorand an der Fakultät für Physik der Universität Karlsruhe (TH)