CWP-549
March 2006

A hybrid formulation of map migration and wave-equation-based migration using curvelets

Huub Douma

— Doctoral Thesis —
Geophysics

Defended on March 31, 2006

Committee Chair:	Prof. John A. Scales
Advisor:	Prof. Roelof K. Snieder
Co-advisor:	Prof. Maarten V. de Hoop
Committee members:	Prof. Gary R. Olhoeft
	Prof. Irina N. Khindanova
	Prof. Kenneth L. Larner

Center for Wave Phenomena
Colorado School of Mines
Golden, Colorado 80401
(1) 303 273-3557

Abstract

The local directions of a wavefront combined with knowledge of the medium velocity reveal the directions of wave propagation. In reflection seismology this property enables one to identify the location where the wavefield reflected through a one-to-one mapping from the data to the image, known as *map migration*. Using building blocks of seismic data that have locally associated directions, such as curvelets, these directions can be used in a map-migration-based procedure to image seismic data. Focussing on sparsely representing the imaging operator and using common-offset (CO) time-migration as an example, I show that to leading-order in angular frequency, horizontal wavenumber, and image location, such migration is equivalent to transforming the coordinates of curvelets combined with amplitude scaling. This transformation is calculated using CO map time-migration equations developed here that use the slopes provided by the curvelet decomposition of the data. The combination of map migration and curvelets unifies the finite-frequency character of curvelets with the ability of map migration to handle caustics.

I extend the CO map time-migration equations to transversely isotropic media with a vertical symmetry axis (VTI) for pure- and converted-mode waves, allowing estimation of the anisotropic parameters in the context of migration velocity analysis. In the context of moveout analysis in common-midpoint (CMP) gathers, I develop a rational-interpolation-based approach that overcomes the bias in the nonhyperbolic moveout equation of Alkhalifah & Tsvankin (1995) for qP-waves, and combines unbiased estimation of the anellipticity parameter η with accurate moveout correction for arbitrary levels of anisotropy and any offset-to-depth ratio of interest. For a single horizontal VTI layer, fixed η, and fixed zero-offset traveltime, the influence of the normal moveout velocity turns out to be limited to a stretch along the offset axis.

Rather than imaging the subsurface using a single-scattering assumption, coda wave interferometry *detects* small changes by cross-correlating multiply scattered wavefields related to an unchanged and changed medium. Because in the presence of noise this cross-correlation is biased, assuming that the underlying stochastic processes of the noise realizations in both signals are mutually uncorrelated and stationary with zero mean, I derive a correction factor that allows removal of this bias.

SAMENVATTING (SUMMARY IN DUTCH)

Als de propagatie snelheid van een medium bekend is, kan de propagatie richting van een golf-front bepaald worden aan de hand van de lokale oriëntatie van het desbetreffende golf-front. In reflectie seismologie bepaalt deze oriëntatie de richting waarin men in de aarde moet kijken om de lokatie van de reflectie te bepalen. Dit principe maakt het mogelijk om deze lokatie te bepalen met behulp van een injectieve afbeelding van de seismische data naar de seismische afbeelding, veelal bekend als *map migratie*. Door de seismische data te representeren als een superpositie van functies met lokaal geassocieerde richtingen, kunnen deze richtingen gebruikt worden in een afbeeldings procedure die is gebaseerd op map migratie.

In dit werk worden curvelets geïdentificeerd als kandidaten voor een simultaan ijle representatie van de seismische data en de afbeeldings operator. Door mij voornamelijk te richten op de ijle representatie van de afbeeldings operator en tijd migratie met gemeenschappelijk epicentrale afstands (GEA) data als voorbeeld te gebruiken, laat ik zien dat tot leidende orde in de hoek-frequentie, het horizontale golf-getal, en de afbeeldings lokatie, zulke migratie equivalent is aan een coordinaten transformatie van de curvelets in de data, in combinatie met een amplitude schaling. Deze transformatie wordt berekend met behulp van map tijd-migratie vergelijkingen voor GEA data, die ik hier afleid. Deze vergelijkingen gebruiken de hellingen die worden bepaald door de curvelet decompositie van de data. Aangezien de data kunnen worden gereconstrueerd met relatief weinig curvelets hoeft deze transformatie voor relatief weinig curvelets te worden uitgevoerd. De nauwkeurigheid van deze transformatie wordt bestudeerd aan de hand van numerieke voorbeelden die bevestigen dat de leidende orde benadering een goede benadering is voor tijd migratie met GEA data. Het gezamenlijke gebruik van map migratie en curvelets combineert het eindige-frequentie karakter van curvelets met het aankunnen van caustieken van map migratie. Aangezien deze karakteristieken normaliter worden toegeschreven aan afbeeldings procedures gebaseerd op de golf-vergelijking, kan dit werk worden gezien als een hybride formulering tussen map migratie en seismische migratie gebaseerd op de golf-vergelijking. De essentiële ingrediënten in de afleiding van de leidende orde benadering voor tijd migratie blijven van toepassing op het geval van diepte migratie. Ik schets hier een mogelijke manier om vervorming van een curvelet bovenop de leidende orde benadering in rekening te brengen, maar benadruk dat deze methode nog geverifiëerd moet worden in de toekomst.

Ik breid de GEA tijd-migratie vergelijkingen voor isotrope media uit naar het geval van transversale isotropie met een verticale symmetrie as (VTI) voor zowel geconverteerde als ongeconverteerde golven, die de schatting van de relevante anisotropie parameters mogelijk maken in de context van migratie snelheids analyse. In de context van move-out analyse in gemeenschappelijke midpunt data (GMP), presenteer ik een methode gebaseerd op rationele interpolatie die de systematische fout in de niet-hyperbolische move-out vergelijking van Alkhalifah & Tsvankin (1995) voor qP-golven voorkomt. Deze methode maakt het mogelijk om de anellipticiteits-parameter η te schatten in combinatie met nauwkeurige move-out

correctie, voor willekeurige anisotropie sterkte en willekeurige afstand-diepte ratio. De methode wordt getest met behulp van zowel numerieke als veld data. Voor een enkele horizontale VTI laag, vaste η, en vaste nul-afstand reistijd, blijkt de invloed van de normale move-out snelheid beperkt te zijn tot het uiteen trekken of in elkaar duwen van de afstands-as in GMP data.

Terwijl seismische migratie typisch gebruik maakt van de aanname dat het golfveld slechts één keer verstrooid is, wordt in *coda golven interferometrie* een kleine verandering in het medium gedetecteerd door gebruik te maken van het meervoudig verstrooide golfveld. In deze methode worden de meervoudig verstrooide golfvelden van een onverstoord en verstoord medium met elkaar gecorreleerd. Wanneer de data ruis bevat, zal deze correlatie in het algemeen een systematische fout bevatten. Daarom leid ik hier een correctiefactor af die het mogelijk maakt te corrigeren voor deze fout, waarbij ik aanneem dat de onderliggende stochastische processen van de ruis realisaties in de signalen van zowel het onverstoorde als verstoorde medium, wederzijds ongecorreleerd, stationair, en gemiddeld nul zijn. De methode wordt getest met behulp van zowel numerieke als veld data.

To my parents, Jan-Dirk and Monica Douma

"It is harmful for the creative work of the mind,
if the intelligence inspects too closely the ideas
already pouring in, as it were, at the gates. Regarded by itself,
an idea may be very trifling and very adventurous,
but it perhaps becomes important on account of the one that follows it;
perhaps in a certain connection with others, which may seem equally absurd,
it is capable of forming a very useful construction.
The intelligence cannot judge all these things
if it does not hold them steadily enough
to see them in connection with others.
In the case of a creative mind, however,
the intelligence has withdrawn its watchers from the gates,
the ideas rush in pell-mell, and it is only then
that the great heap is looked over and critically examined."

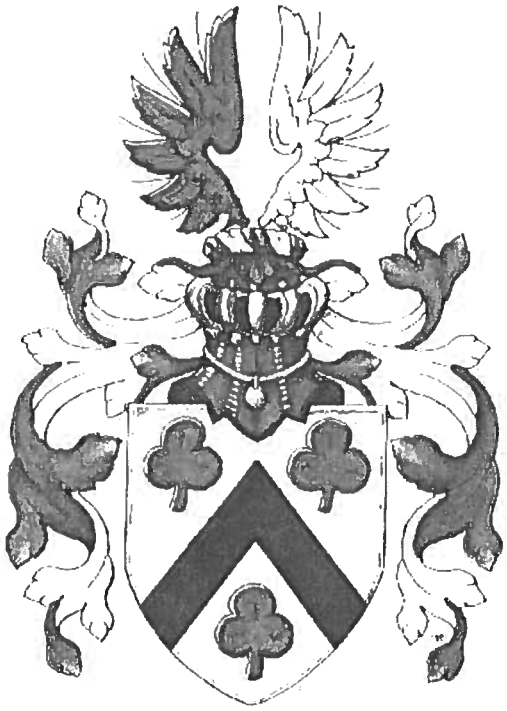
— Sigmund Freud
The Interpretation of Dreams

"Time ... time won't leave me as I am
but time, won't take the boy out of this man"

— Bono/U2
City of Blinding Lights

"The unconscious sends all sorts of vapors, odd beings, terrors,
and deluding images up into the mind ... for the human kingdom,
beneath the floor of the comparatively neat little dwelling that
we call our consciousness, goes down into unsuspected Alladin caves.
There not only jewels but also dangerous jinn abide: the
inconvenient or resisted psychological powers that we have not thought
or dared to integrate into our lives. ... These are dangerous because
they threaten the fabric of the security into which we have built
ourselves and our family. But they are fiendishly fascinating too,
for they carry keys that open the whole realm of
the desired and feared adventure of the discovery of the self."

— Joseph Campbell
The Hero with a Thousand Faces



Douma

Table of Contents

Acknowledgments

In the past few years I have been given the opportunity to fulfill a long cherished dream to acquire a PhD-degree. This opportunity was given to me by my advisor Dr. Roel Snieder, who has been a great inspiration ever since I did my M.Sc. thesis work under his supervision at Utrecht University. Therefore I first want to express my deep gratitude to you, Roel. It has been a privilege, and at the same time a pleasure, to be able to interact with you during these years, and to learn from the many insightful discussions on both science and life in general. I thank you for your guidance and inspiration, for putting your trust in me and standing by me in difficult times, but most of all for your selfless and subtle ways to encourage me to try to find my own direction as both a researcher and a person. Besides your lasting impact on me as a pragmatist and as a versatile and creative researcher, the expression on your face while going down the slide in Waterworld, the jokes from van Kooten en de Bie ("Hij zegt gewoon héele goeie dinguh ...") and de mannen van de radio ("Oprotten! De plant."), as well as the many laughs near "your other office with the comfortable seats", will remain with me forever!

The collaboration with my co-advisor Dr. Maarten de Hoop has been a captivating learning experience in many different ways. Maarten, I want to thank you for your never-ending drive and contagious enthusiasm when it comes to science in general, for pointing out the existence of the curvelet frame, and for making me understand the importance of a carefully chosen research topic. Your passion for mathematics will likely prove to be a lifelong inspiration.

I would like to specifically mention the influence Dr. Ken Larner has had on me during this journey. His methodistic ways to force one to "say what you mean" in both oral and written communication, and his thorough and critical questioning, are values that will stay with me for the rest of my life. Ken, your lessons in dealing with "politically loaded" situations and correspondence will likely prove invaluable in my future career. I feel privileged to have been able to do my PhD-studies at the Center for Wave Phenomena (CWP) while it was under the spell of your subtle yet strong leadership.

Also, I want to thank Emmanuel Candés for providing me with an early version of the digital implementation of the curvelet transform in 2004 that allowed me to start the research on seismic imaging with curvelets, which resulted in Chapter 4 of this thesis. In addition, I want to thank Dr. John Scales, Dr. Gary Olhoeft, Dr. Ken Larner, and Dr. Irina Khindanova for serving as members on my thesis committee and for their continued support over the past few years. The helpful discussions regarding anisotropy with Dr. Ilya Tsvankin are much appreciated.

Studying at CWP has been both a pleasure and a great inspiration. I want to thank everybody at CWP for their help and support, especially Michelle Szobody for always accurately taking care of all the necessary paperwork with a smile on her face, and Barbara

McLenon for her support with manuscript submissions and publications, as well as her loving way in showing interest in one's personal life. From the students I would like to single out Matt Haney for providing some late-night musical extravaganzas and for proving that science and rock-'n-roll can form a succesful symbiosis, Kurang Mehta for showing me the power of pragmatism, Matt Reynolds for being a partner in crime in the dark hours of real and functional analysis, Yaping Zhu for being so resourceful and creative in scientific discussions, and Ivan Vasconcelos for making me remember the power of confidence.

I want to thank my former employer, Western Geophysical (now WesternGeco), and in particular Dr. Swavek Deregowski and Peter Nuttall, for providing financial support during the first half of my studies that helped ease the financial burden from going from a well-paid job back to a student budget. Also, I would like to thank Dr. Robert Bloor (formerly Western Geophysical, now GX Technology) for making me aware of map migration, which gave me the idea to do part of the work reported in this thesis. Thanks go to the guys at GX Technology, especially Alexander Calvert and Edward Jenner, for their help in testing the rational interpolation based nonhyperbolic moveout analysis on real data.

From my friends I want to specifically mention Dr. Jeroen van Hunen, who coincidently managed to find a postdoctoral research position in Boulder a year after I started living there. Jeroen, thanks for bringing back some of the good old times from Utrecht to Boulder, for doing the unthinkable of leaving me in the dust while running (yes Erik, it's true), for the biking session with the infamous Mr. Salt Face (alias Erik), and dumping your whole inventory with us when leaving (now I'm the proud owner of the orange glass plates, Jelle). Holding your few-hours-old daughter in the hospital in Denver was a highlight! Nick, thanks for picking me up from the street and riding me back and forth to Golden many times, for reintroducing me to swimming and encouraging me to join the masters-swimming program, and for the long-winded and absent-minded phone messages. Jelle, Jeroen, Erik, Sander, Martijn en Pieter, thanks for being there on the other end of the line in the dark moments when I was lost and longing to be back home in the Netherlands. Also, many thanks to Rembrandt Zuijderhoudt for the late-night altruistic tele-coaching sessions when I could hardly see the light; I hope our paths will cross again.

I thank my cousin Dr. Jan Douma for the many stimulating talks about geophysics, especially during my high-school years when he drove to visit us in that magnificent run-down orange Opel Kadett, causing me to choose a career in geophysics. Jan, if it wasn't for you and that impressive car, I might have strayed to become a beachbum.

At the time when I was playing with the idea to start a PhD degree at CWP, I was working in London, where I had met this beautiful Polish woman named Eva. When I asked her if she would want to join me in this adventure, she wholeheartedly answered "Yes", and we subsequently got married in that magical castle in Scotland. Eva, I want to thank you for your generosity, your kind-hearted nature, your confidence in me, and for standing by my side through the difficult times in this journey. Without your mental and emotional support, and your loving ways in life, I could not have done this. Thank you!

Finally, I dedicate this thesis to my parents. It is your loving support and your encouragement throughout my whole life that have brought me to where I am today, doctor of philosophy in geophysics. Without all the opportunities that you have given me, for

which I now realize you had to work so hard, without the many stimulating words and conversations, as well as your unconditional confidence in me, I would not have been able to achieve this. Really, I owe everything to you. Therefore father, in your honour, I finish this thesis by saying "Fryslân boppe!"

Boulder, March 2006

Chapter 1

Introduction and outline

In a surface seismic experiment, seismic waves that have traveled through the earth and originated at its surface are recorded at the surface. Because of the bandlimited nature of the seismic sources, combined with the anelasticity and the dispersive nature of the earth's interior, the recorded wavefield is a superposition of bandlimited wavefronts that were partially reflected and transmitted at discontinuities such as geological interfaces and faults in the earth. Since a wavefront can be characterized by its location, locally associated directions, and frequency content, the nature of the recorded data can be characterized by bandlimitation and locally associated directions.

In seismic data processing, the recorded wavefront is represented by discrete sample values every so many milliseconds and every so many meters; that is, we are using delta functions as the building blocks of seismic data. Since the Fourier transform of a delta function is unity, a delta function is characterized by infinite bandwidth and all directions. Therefore, the current building blocks of seismic data do not encompass the physical nature of the data, i.e., bandlimitation and locally associated directions. So the natural question arises whether we can find building blocks of seismic data that have the character of the data already built into them, i.e., they are bandlimited, local, and have locally associated directions. It seems intuitive that such basis functions allow a sparser representation of the data since they encompass the character of the data. Moreover, such basis functions could potentially also allow for a sparsification of seismic imaging because of their locally associated directions. To understand this, I first review the essence of seismic imaging methods that are based on the high-frequency approximation.

Seismic imaging algorithms that are based on high-frequency asymptotics, such as Kirchhoff migration which makes use of ray theory, involve integration over diffraction surfaces that connect sample values in the data that would lie on the recorded wavefront if a wavefield were scattered once at a particular point in the (modeled) subsurface. This integration is a many-to-one mapping. Because such integration needs to be done for every point in the modeled subsurface to build up an image, and because for each of these points the associated diffraction surfaces need to be calculated based on the subsurface model, it is costly computationally. In this method the slopes of reflections in the data are used implicitly only.

In the high-frequency approximation, seismic waves travel along rays through the subsurface. The derivative of the traveltime with respect to surface location, usually referred to as the slope of a reflection in the data, together with the velocity, determine the direction normal to the wavefront recorded at the surface. Therefore, calculating the slopes at source

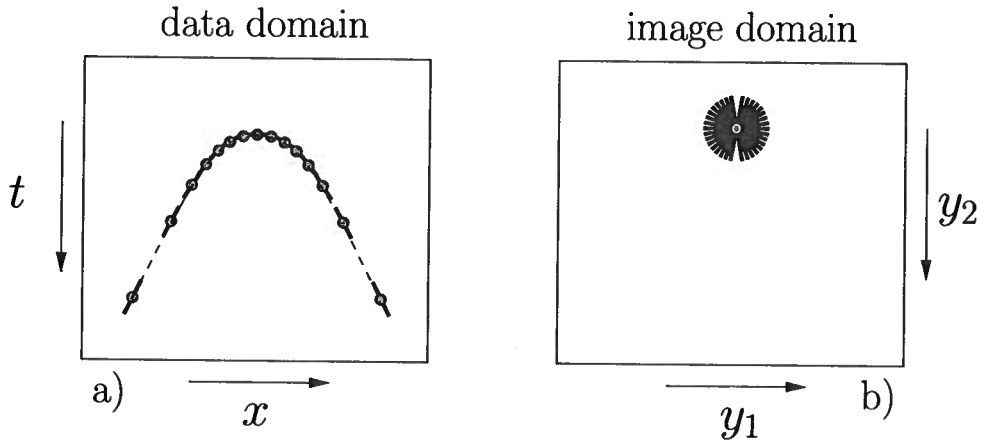


Figure 1.1. a) Line elements associated with a diffraction surface in the data, and b) the corresponding line elements in the image domain. Map migration maps each line element in the data domain, i.e., a location, traveltime, and slope, one-to-one to a line-element in the image domain, i.e., a horizontal and vertical location and a certain orientation. This illustration pertains to post-stack migration, i.e., migration of data for coincident source and receiver.

and receiver locations (i.e., in a common-receiver and common-source gather, respectively), determines the "take-off" directions of the associated rays from these locations. Therefore, once the slopes at the source and receiver locations are given (together with the velocity), the location of the reflector in the earth can in principle be found using ray-tracing; i.e., the crossing-point of the rays from the source and receiver determines the imaged location of the reflector [e.g, Sword (1987)]. If the velocity model is accurate, the resulting traveltime from the source to the reflector and back to the receiver equals the measured traveltime. With use of Snell's law at the imaged reflector location, the orientation of the reflector can also be established. Therefore, the mapping from the source and receiver location, the traveltime, *and* the local slopes at the source and receiver, to the imaged reflector location *and* orientation, is one-to-one (see Figure 1.1). This mapping is referred to as pre-stack *map migration*. The combination of location, traveltime, and orientations constitutes an element of *phase-space*. Map migration is therefore a mapping from the phase-space of the data to the phase-space related to the image.

This explanation of the one-to-one mapping property of map migration can be made more intuitive by considering the following thought-experiment. Suppose a young woman is having a picnic in the woods with her lover, Vincent van Gogh. As they are enjoying each other's company and the food, they hear the howling of a nearby hungry wolf. Poor Vincent, who has only one ear, has no way of identifying the direction from which the sound came. As far as he is concerned, the wolf could be located anywhere on a circle around

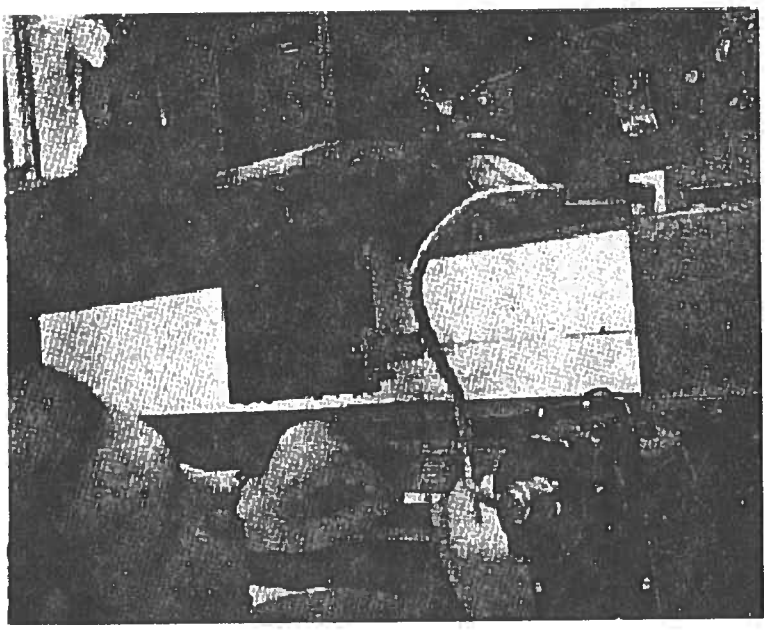


Figure 1.2. Ken Larner (with hair) messing with the "sonograph" in the previous century. Figure taken from Rieber (1937).

him with radius equal the product of the traveltime and the speed of sound. This situation is analogous to the many-to-one mapping in Kirchhoff migration. Vincent's young love, however, picks up a traveltime difference between one ear and the other, and, by turning her head, she can find the direction when both ears hear the wolf at the same time. Looking in that direction she either sees the wolf straight in front of her or the wolf is straight behind her, in which case she'd better turn around before the wolf alters the course of history and robs Vincent of his other ear. (In seismic imaging we can ignore this ambiguity because we know that all energy was reflected below the earth's surface.) Therefore, the traveltime difference, i.e., the traveltime derivative with respect to the location of the ear, allows the young woman to identify the direction from the cry of the wolf. This situation is analogous to zero-offset (or post-stack) map migration in the context of the exploding reflector model (Claerbout, 1985), and thus to the one-to-one mapping property of map migration. Note that the ability to identify a traveltime difference is limited by the size of the wavelength relative to the separation of the ears. That is, if the wavelength is much longer than the distance between both ears, both ears hear essentially no traveltime difference, and one is no longer able to identify the direction to the source of the sound. This is why it does not matter where the whoofer of a sound system is placed in the room, as it emits the low bass frequencies that have wavelengths much larger than the distance between our ears (if for your sound system it does matter where you place your whoofer, you'd better start looking for your receipt). Therefore the one-to-one property of map migration is actually the result of the high-frequency asymptotic approximation. The importance of slopes and

(controlled) directional analysis was already realized by Rieber (1936, 1937), who devised an apparatus called "the sonograph" to perform local slant-stacks through optical and electrical adjustments of the recorded data (see Figure 1.2). This device allowed the determination of the local slopes in the data prior to the digital age.

There exist special cases when the mapping from the data phase-space to the image phase-space, i.e., map migration, is not one-to-one [see Stolk (2000) or de Hoop & Brandsberg-Dahl (2000) for some examples]. In these cases it can for example happen that a combination of a traveltime, the source and receiver locations, and the associated local slopes can be mapped to different reflector locations. Thus, this phase-space mapping is generally not one-to-one (Guillemin, 1985), as mentioned in Chapter 2 of this thesis. In practice, however, these special cases are ignored because they cannot be identified with practical acquisition geometries. For all practical purposes and for the remainder of this introduction, I therefore tacitly assume that these special cases do not occur. Hence the mapping is simply considered to be one-to-one.

Knowing the crucial role of slopes in seismic imaging, it seems logical to use basis functions for seismic data that have locally associated directions. Then a simple projection of the data onto this basis provides the local slopes in the data that can subsequently be used in a pre-stack (amplitude-preserving) map-migration procedure to image the data. This allows the many-to-one mapping of integration along diffraction surfaces to be replaced by the one-to-one mapping provided by map migration. In other words, such functions allow a sparse representation of the imaging operator, and thus a potential speed-up of current imaging algorithms[1].

Knowing that seismic data can be characterized by bandlimitation and locally associated directions, in seismic imaging the use of basis functions that naturally encompass this character, could have a double potential gain. First, as explained above, it seems intuitively plausible that such basis functions would allow for a sparse representation of the data because the character of the data is already built into the basis elements. In addition, using only the basis elements that contribute significantly in the data representation (i.e., the elements that have large coefficients after projection of the data onto the basis), their locally associated directions could be used with the one-to-one mapping provided by map migration to image the data. Then, not only would we have fewer basis elements to work with, but each element would be subjected also to a one-to-one instead of a many-to-one mapping. This is the main underlying motivation of this thesis. I concern myself with the sparse representation of the imaging operator only. The sparse representation of the data is not studied in-depth, i.e., I do not investigate any sophisticated denoising methods to achieve a sparse representation of the data.

[1] I conceived this idea when working at Western Geophysical (now WesternGeco) as a research scientist, when map migration was brought to my attention by Dr. Robert Bloor.

Outline

In Chapter 2, I explain the one-to-one mapping property from the phase-space of the data to the phase-space related to the image, and make clear its connection with map migration. I exemplify this one-to-one mapping property by deriving closed-form 3D pre-stack map time-migration expressions for qP-qP, qP-qSV, and qSV-qSV waves in homogeneous transversely isotropic (TI) media with a vertical symmetry axis (VTI), which specialize to the expressions for P-P, P-S, and S-S waves in homogeneous isotropic media. These expressions can in principle be used to determine the anellipticity parameter η (which quantifies the deviation of the slowness surface from an ellipse) and the zero-dip normal-moveout (NMO) velocity V_{NMO} for such media in the context of time-migration velocity analysis. In addition, I present closed-form expressions for pre-stack map time-migration and demigration in the common-offset domain, for pure-mode (P-P or S-S) waves in homogeneous isotropic media, that use only the slope in the common-offset domain. This provides an additional advantage over methods that use both slopes in the common-offset and common-midpoint domain, especially since estimating slopes can be cumbersome in the presence of noise. Considering kinematics only, Chapter 2 explains that map migration (or demigration) provides a suitable vehicle to sparsely represent the imaging (or modeling) operator.

The pre-stack map time-migration equations for VTI media presented in Chapter 2 are valid for arbitrary levels of anisotropy. Many rocks, however, exhibit only weak anisotropy (Thomsen, 1986). I present the simplifying expressions for pre-stack and post-stack map migration in homogeneous weakly VTI media in Chapter 3. The resulting expressions are explicit in the relevant parameters η and V_{NMO}.

In Chapter 4, curvelets are identified as suitable candidates for the basis functions[2], mentioned earlier in this introduction, that would allow a simultaneous sparse representation of the seismic data and the imaging operator. Roughly speaking, curvelets are pieces of bandlimited plane waves[3] that have associated dominant local directions. In Chapter 4 I concern myself mainly with the sparse representation of the imaging operator[4], and I present a leading-order sparse representation of the imaging operator based on map migration, in the context of common-offset time migration. This representation allows the leading-order contribution of such imaging to be done through a simple transformation of coordinates of each input curvelet (resulting from the curvelet decomposition of the data), calculated

[2]Curvelets actually do not constitute a basis but rather a *tight frame* for functions in L^2 (see Appendix I for an explanation of tight frames). Omitting most details that are beyond the scope of this introduction, it suffices to understand that curvelets allow a complete reconstruction of the data. This means that if all curvelets are used, the data can be reconstructed without any residual. To avoid distraction of the main message of this introduction by including too many details, I therefore simply refer to curvelets as basis functions, even though they are really not.

[3]The difference between this rough description and the true character of curvelets is that a bandlimited plane wave has associated with it only one $\boldsymbol{k}$-direction, while a curvelet is associated with a small range of $\boldsymbol{k}$ vectors.

[4]The sparse representation of the data is obtained simply by using only the largest coefficients to reconstruct the data. These coefficients are determined based on a direct comparison of the reconstructed data with the original data for varying threshold percentages. Throughout this thesis I do not study any sophisticated denoising algorithms to obtain a sparse representation of the data.

using map migration, combined with amplitude scaling. Numerical verification of this approximation shows that the leading-order approximation to common-offset (CO) time-migration provides a good approximation to such imaging. This highlights the potential of using curvelets in seismic imaging. Note that the combination of the finite-frequency character of curvelets with map migration in essence combines high-frequency asymptotic-based Kirchhoff imaging with finite-frequency-based methods such as wave-equation-based imaging.

The derivation of the leading-order sparse representation for CO time-migration presented in Chapter 4 can be extended to general heterogeneous anisotropic media. The essence of the presented derivation — a stationary phase evaluation of an oscillatory integral, combined with the linearization of the remaining phase and the linearization of the image around the map migrated location (here shown in the context of time-migration) — continues to hold for depth migration. For that case, it remains to be seen how accurate the leading-order approximation is. I explain that it is my current understanding that deformation of the curvelet beyond the leading-order approximation can be done using an approach similar to the one outlined by Smith (1998), although this idea remains to be verified in the future. The computational cost of correction for the deformation of curvelets beyond the leading-order approximation is currently unknown. Therefore, the question as to whether a curvelet representation of the imaging operator allows for a significant speed-up of seismic imaging remains currently open. The benefit of using curvelets in seismic imaging, however, is not limited to a potential gain in computational efficiency. The bandlimited nature of curvelets allows migration velocity analysis to be done as a function of frequency. Such analysis would be physically relevant as it is known that a wave averages the medium properties at a scale related to the wavelength upon propagation [the first Fresnel zone is proportional to $\sqrt{\lambda}$ with λ the wavelength (e.g., Kravtsov (1988); Spetzler & Snieder (2004)]. Hence, waves of different frequency indeed observe the earth at different scales. Using curvelets in seismic imaging could allow naturally for such analysis. Although for constant-velocity media the medium is the same at every scale, I show in Chapter 4 that, in principle, images can be made for different frequencies by simply choosing to use curvelets that have certain frequencies only. This possibility is a straightforward consequence of an appropriate representation of both the seismic data and the imaging operator. Moreover, the locally-associated main directions of curvelets can be used for controlled-illumination (Rietveld & Berkhout, 1992). Considering the current interest in sub-salt imaging, controlled illumination studies can have important consequences for acquisition design, by establishing the image quality as a function of illumination direction. With curvelets, controlled illumination can be achieved by simply choosing to use curvelets with certain main directions only. Again, this opportunity is a straightforward consequence of an appropriate representation of the data and the imaging operator; no additional preprocessing of the data is required. In Chapter 4 I illustrate this in the context of time migration.

The pre-stack map time-migration equations for qP-waves presented in Chapter 2 can be used to estimate the anellipticity parameter η by optimizing a misfit function that measures the differences between an imaged reflector (position, time, *and* slope) for different offsets by letting η vary. Even though I do not demonstrate such an approach, the presented

equations allow the estimation of η to be put in the context of migration velocity analysis. Conventionally, initial estimates of η are obtained by analyzing the moveout of qP-waves in common-midpoint (CMP) gathers and ascribing the deviation from hyperbolic moveout to the presence of anisotropy. In practical seismic data processing, such analysis is often still based on the nonhyperbolic moveout equation of Alkhalifah & Tsvankin (1995). In Chapter 5, I show that using this equation for this analysis induces a bias in the estimated value of η, which increases as η becomes larger (roughly speaking, this bias becomes noticeable when $\eta > 0.1$). More so, for larger offset-to-depth ratios, the Alkhalifah-Tsvankin equation fails to accurately flatten the moveout in a CMP gather even with the biased values of η. To overcome these problems, I present a method based on rational interpolation that combines unbiased estimation of η with accurate moveout correction for arbitrarily large offset-to-depth ratios. Analysis of the dependence of moveout on η and V_{NMO} reveals that, under the generally accepted assumptions that the traveltimes of qP-waves in TI media depend mainly on η and V_{NMO} (i.e., we can set the Thomsen parameter $\delta = 0$ [5]), the influence of V_{NMO} on the moveout curve is limited to a stretch of this curve along the offset axis. This observation allows the traveltimes needed for the rational interpolation to be found from a (small once-in-a-lifetime) precalculated table that holds the traveltimes as a function of η for a number of different offset-to-depth ratios, along with a chosen reference value of V_{NMO} (typically $V_{NMO} = 1$ km/s). This renders the method highly efficient (i.e., as efficient as using expressions describing the moveout explicitly in terms of η and V_{NMO}). A field data example confirms the findings from several numerical tests.

The rational-interpolation-based method for nonhyperbolic moveout analysis presented in Chapter 5 is just like the nonhyperbolic moveout equation in Alkhalifah & Tsvankin (1995), based on a single horizontal VTI layer. Therefore, this method allows the estimation of *effective* values of η for an overburden treated as if it consisted of a single effective layer. As pointed out by Winterstein (1986), however, in order for the estimated anisotropy parameters to be used as lithology indicators, it is crucial to obtain *interval* estimates of those parameters. Therefore, in Chapter 6, I extend the rational-interpolation-based method for nonhyperbolic moveout analysis of qP-waves in VTI media to horizontally layered VTI media, to allow the estimation of the interval values of η and V_{NMO}. The method is based on the $\tau - p$ curve, which represents a plane-wave decomposition. This is the natural decomposition for horizontally layered media because cusps do not occur in this domain and the horizontal slowness is maintained upon propagation through such media. Based on the acoustic approximation of Alkhalifah (1998), I derive expressions for the two-way traveltime t and the associated offset x as a function of the horizontal slowness p, which are explicit in the interval values of η and V_{NMO}. These expressions are used to calculate the support points needed for a rational interpolation. The efficiency of the method stems from the fact that highly accurate traveltime curves are obtained with only few support points. The accuracy of the method is illustrated with a synthetic data example. Because the single layer case is simply a special case of multi-layered ones, this method overcomes

[5]The Thomsen parameter δ determines the angular dependence of the qP-wave phase velocity in the vicinity of the symmetry axis, which is vertical in VTI media.

the need for the tabulated approach as presented in Chapter 5 for a single horizontal VTI layer.

As mentioned previously, the use of curvelets in seismic imaging essentially combines high-frequency asymptotic-based Kirchhoff imaging with finite-frequency wave-equation based imaging. Typically finite-frequency-based methods make use of cross-correlation in one way or another [e.g., Claerbout's imaging principle is based on cross-correlation of the up- and down-going wavefields (Claerbout, 1971), while wave-equation-based traveltime tomography (Luo & Schuster, 1991; Marquering *et al.*, 1999) uses cross-correlation to estimate the traveltime differences between observed and synthetic seismograms]. Coda wave interferometry (Snieder *et al.*, 2002) is a finite-frequency method wherein cross-correlation is used to detect and infer small changes in a multiply scattering medium, rather than imaging them based on a single-scattering assumption. In this method the cross-correlation between two seismograms, one related to an unperturbed medium and the other to the perturbed medium, is mapped into an estimate of the change in the medium. If both seismograms contain noise, the cross-correlation will be changed, thus introducing a bias in the estimate of the change in the medium. In Chapter 7, I present a correction factor that allows the removal of bias in the cross-correlation, under the assumptions that the stochastic processes underlying the noise realizations in both signals are mutually uncorrelated and stationary with zero mean. Because the medium is assumed not to change over the duration of the experiment, different (non-overlapping) time windows in the coda allow for independent estimates of the medium perturbation. If the seismograms are contaminated with noise, only those time windows should be used for which the amplitude of the coda is significantly above the ambient noise level. This limits the usable number of independent time windows. The proposed correction factor thus allows more time windows further into the decaying coda to be used, and hence allows for a reduction of the error bars on the medium perturbation estimates.

Chapter 2

Explicit expressions for pre-stack map time-migration in isotropic and VTI media and the applicability of map depth-migration in heterogeneous anisotropic media[1]

2.1 Summary

We present 3D pre-stack map time-migration in closed form for qP, qSV and mode-converted waves in homogeneous transversely isotropic (TI) media with a vertical symmetry axis (VTI). As far as pre-stack time-demigration is concerned, we present closed-form expressions for the mapping in homogeneous isotropic media, while for homogeneous VTI media we present a system of four nonlinear equations with four unknowns that needs to be solved numerically. The expressions for pre-stack map time-migration in VTI homogeneous media are directly applicable to the problem of anisotropic parameter estimation (i.e., the anellipticity parameter η) in the context of time-migration velocity analysis. In addition, we present closed-form expressions for both pre-stack map time-migration and demigration in the common-offset domain for pure mode (P-P or S-S) waves in homogeneous isotropic media, that use only the slope in the common-offset domain as opposed to slopes in both the common-shot and common-receiver (or equivalently the common-offset and common-midpoint) domains. All time-migration and demigration equations presented can be used in media with mild lateral and vertical velocity variations, provided the velocity is replaced with the local RMS velocity. Finally, we discuss the condition for applicability of pre-stack map depth-migration and demigration in heterogeneous anisotropic media that allow the formation of caustics, and explain that this condition is satisfied if, given a velocity model and acquisition geometry, one can map depth-migrate without ambiguity in either the migrated location or orientation of reflectors in the image.

2.2 Introduction

The geometry of seismic migration can be understood and described in terms of surfaces of equal traveltime, i.e., isochrons. Migration encompasses the integration of signal

[1]This Chapter, along with appendices A and B, has been published as: Douma, H. and de Hoop, M. V. 2006. Explicit expressions for pre-stack map time-migration in isotropic and VTI media and the applicability of map depth-migration in heterogeneous anisotropic media, Geophysics **71**, no. 1, p. S13-S28.

processed data along diffraction surfaces related to these isochrons. In terms of linear filter theory, in homogeneous media, the image is a convolution of the impulse response, shaped in accordance with the isochrons of the migration operator, with the data. This approach uses the positions, traveltimes, and amplitudes of the events in the data, and thus uses the information given by the reflection slopes implicitly.

In the high-frequency approximation, seismic waves (or wave singularities) propagate along rays through the subsurface. Provided the velocity in the earth is known, reflection slopes in the data determine the directions of such rays at the recording surface (or the singular direction of the wavefront set of the recorded wavefield). Therefore, once the traveltime and the slopes at the source and receiver are known along with the velocity, the location and local dip of a reflector in the subsurface (or image singularity) can in principle be determined with the aid of numerical ray tracing (Červený, 2001).

The determination of the reflector position and orientation from the times and slopes of zero-offset reflection seismic data, is generally referred to as *map migration* (Kleyn, 1977). We refer to such determination from times and slopes of reflection data at the source and receiver locations as pre-stack map migration. In a mathematical context, provided the velocity is known and the medium does not allow different reflectors to have identical surface seismic measurements that persist under small perturbations of the reflectors, the use of the slope information results in a one-to-one mapping from the unmigrated quantities associated with a reflection in the data (given a scattering angle and azimuth), to the migrated quantities associated with a reflector in the image. Collecting the migrated and unmigrated quantities in a 'table', leads to the notion of *canonical relation.*

Here we aim to elucidate that pre-stack map depth-migration and demigration are closely related to the canonical relation of the single scattering imaging or modeling operators in complex media. Imaging artefacts (or 'imaging phantoms') are avoided if the projection of this canonical relation on the unmigrated quantities is one-to-one. This condition was introduced by Guillemin (1985) and further exploited by ten Kroode *et al.* (1998), de Hoop & Brandsberg-Dahl (2000) and Stolk & De Hoop (2002b). We explain this, and make clear that this condition is precisely the applicability condition that allows map depth-migration in inhomogeneous anisotropic media (i.e., in the presence of caustics).

The concept of map migration is certainly not new. Weber (1955) gives an early account of map migration, wherein the zero-offset 3D map migration equations are derived for a constant-velocity medium and arbitrary recording surface. Independently, Graeser *et al.* (1957) and Haas & Viallix (1976) use the slopes in the data to derive the position of a reflector in 3D from zero-offset data for a homogeneous isotropic medium. In an early attempt at the use of numerical ray tracing, Musgrave (1961) uses the slope information to calculate wavefront charts and migration-table lists, and Sattlegger (1964) derives a series expansion for the coordinates of the raypath that uses the slopes in the data. Both methods assume vertically-varying velocity media and can be used for 3D migration of zero-offset data. Reilly (1991) and Whitcombe & Carroll (1994) present successful applications of post-stack map migration on field data.

Map migration has been used for velocity estimation in several different approaches [e.g., Gjoystdal & Ursin (1981), Gray & Golden (1983), and Maher *et al.* (1987)]. To

improve horizon-based velocity model building, map migration has been used in seismic event-picking schemes consisting of map migration, followed by picking of events and slopes, map demigration, and remigration in the updated velocity model. The initial map-migration step in such a scheme attempts to reduce mispositioning of the velocity picks. The idea to use map migration for velocity analysis stems from the sensitivity of pre-stack map migration to the migration velocity model, as pointed out by Sattlegger *et al.* (1980). Sword (1987, p.22) develops a controlled directional reception (CDR) tomographic inversion technique, first suggested by Harlan & Burridge (1983), to find interval velocities from pre-stack seismic data. In that method the slopes (or horizontal slownesses) are picked automatically using the CDR picking technique — slant-stack over a short range of offsets with subsequent picking — developed in the former Soviet Union [e.g., Zavalishin (1981) and Riabinkin (1991)]; but first introduced by Rieber (1936) and later reintroduced by Hermont (1979). Subsequently the estimates of the ray parameters are used to trace rays through the initial estimate of the velocity model, and a depth is found wherein the sum of the traveltime along the downgoing (source) and upgoing (receiver) rays equals the observed traveltime. Then, at this depth, the horizontal distance between the end-points of two rays in the subsurface is minimized using a modified Gauss-Newton method to yield the velocity model.

Recently Iversen & Gjoystdal (1996) performed 2D map migration in arbitrarily complex media using a layer-stripping approach similar to that of Gray & Golden (1983) to achieve simultaneous inversion of velocity and reflector structure; they later extended this method to 2D anisotropic media (Iversen *et al.*, 2000). Their linearized inversion scheme, which minimizes the projected difference along the reflector normal between events from different offsets, uses derivatives of reflection-point coordinates with respect to model parameters as introduced by van Trier (1990), rather than derivatives of traveltimes with respect to model parameters as used in classical tomographic inversion [e.g., Bishop *et al.* (1985)]. Such an approach allows for more consistent event picking because reflectors can be identified in a geological structure. In addition, an initial imaging step generally improves the signal-to-noise ratio allowing for more accurate event picking. Finally, Billette & Lambaré (1998) most recently reiterated the importance of slope information in velocity model estimation while validating that precision in measured slopes, traveltimes, and positions in seismic reflection data is sufficient to recover velocities using stereotomography.

In this paper we first develop closed-form expressions for the geometry of pre-stack map time-migration and demigration in 3D for a homogeneous isotropic medium. These migration equations assume, in addition to the velocity (as is common in seismic imaging), that only the location and the slopes in unmigrated common-offset gathers are known. This is in contrast to the migration equations in 2D presented by Sword (1987, p.22), which also require the slope within the common-midpoint gathers (or, alternatively, the slopes in the common-source and common-receiver gathers). Time migration, which uses the assumption of a root-mean-square (RMS) velocity remains in use. In this context our expressions have current applicability, provided the constant velocity in them is replaced by the local RMS velocity. To complement the work of Alkhalifah & Tsvankin (1995) and Alkhalifah (1996) on velocity analysis in transversely isotropic media, we derive closed-form expressions for 3D pre-stack map time-migration for qP waves, qP-qSV mode-converted waves, and qSV waves

in transversely isotropic (TI) media with a vertical symmetry axis (VTI). These expressions are directly applicable to the problem of anisotropic parameter estimation (such as the anellipticity parameter η) in the context of time-migration velocity analysis. For 3D pre-stack map time-demigration in such media, we present a system of four nonlinear equations with four unknowns, that needs to be solved numerically. Since in the vertical symmetry plane of TI media with a tilted symmetry axis (TTI), the phase and group directions are both in this plane, our results apply also to the vertical symmetry plane of homogeneous TTI media. Furthermore, the kinematic equivalence of TI media and orthorhombic media in the symmetry planes, generalizes our results to these planes in orthorhombic media. We then proceed to explain the applicability of map depth-migration in heterogeneous anisotropic media, and revisit pre-stack map time-migration in homogeneous isotropic media to show that the pre-stack map time-migration and demigration equations *define* the canonical relation of the single scattering modeling and imaging operators in such media. For practical issues such as slope estimation, accuracy and stability of the algorithm, and sampling and grid distortion, we refer to the existing literature [e.g., Kleyn (1977) and Maher & Hadley (1985)].

2.3 Map time-migration and demigration in isotropic media

To illustrate the concept of map migration, we derive explicit 3D pre-stack map migration and demigration equations for isotropic homogeneous media. We assume that preprocessing has already compensated data for any topography on the acquisition surface, and deal with only the geometry of migration. The results, however, could be extended to take geometrical spreading effects into account. For media with mild lateral and vertical velocity variations, these equations can be used provided the velocity is replaced with the local RMS velocity.

For 2D pre-stack map time migration, Sword (1987, p.22) derived closed-form expressions for the migrated location and reflector dip using the horizontal slownesses at both the source and the receiver. It turns out, however, that for pre-stack map time-migration and demigration in homogeneous isotropic media, we do not need the angles at the source and the receiver given by the horizontal slownesses. We demonstrate this by deriving closed-form expressions for pre-stack map time migration and demigration in 3D, that use only the slopes in common-offset gathers. These expressions provide a practical advantage over existing closed-form solutions that use both slownesses, since only one slope needs to be measured instead of two. Such a reduction, unfortunately, no longer holds in heterogeneous or anisotropic media.

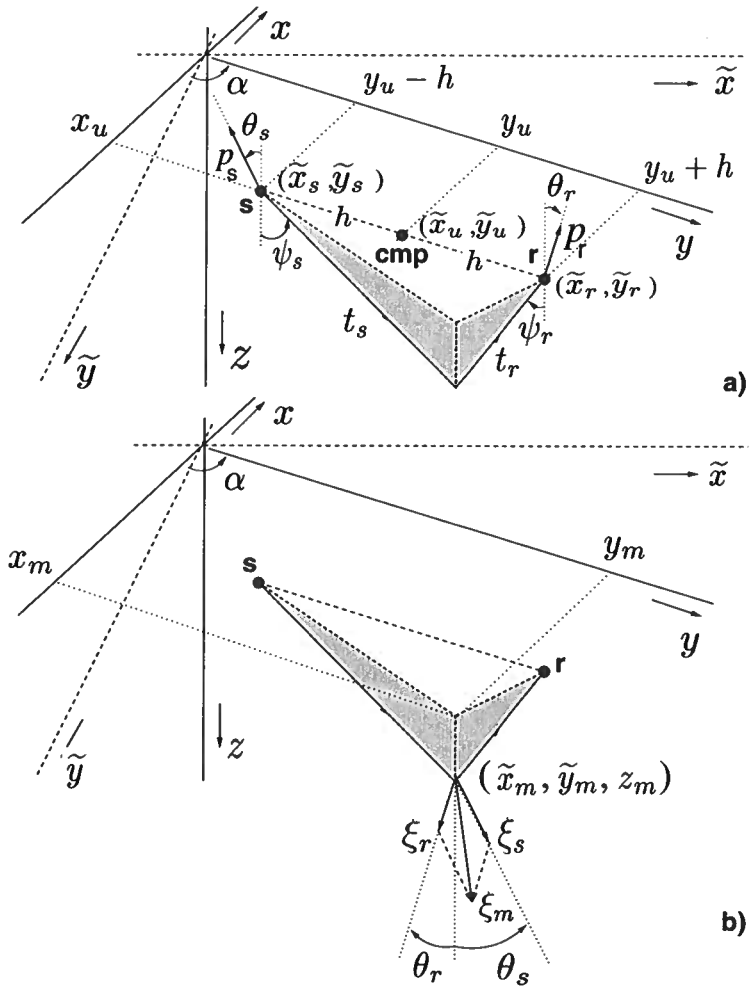


Figure 2.1. Geometry defining the DSR equation for a transversely isotropic homogeneous medium with a vertical symmetry axis, in both the unrotated $(\tilde{x}, \tilde{y})$ and rotated (x, y) reference frames. a) Coordinates describing the data: source location $(\tilde{x}_s, \tilde{y}_s)$ or $(x_s = x_u, y_s = y_u - h)$ with h the half-offset, receiver location $(\tilde{x}_r, \tilde{y}_r)$ or $(x_r = x_u, y_r = y_u + h)$, midpoint location $(\tilde{x}_u, \tilde{y}_u)$ or (x_u, y_u), two-way traveltime $t_u = t_s + t_r$, and the slowness vectors at the source $(\boldsymbol{p}_s)$ and receiver $(\boldsymbol{p}_r)$; $\theta_{s,r}$ and $\psi_{s,r}$ are, respectively, the phase and group angles with the vertical symmetry axis at the source (s) and receiver (r). The angle α is the angle between the $\tilde{y}$ and y direction, i.e., the acquisition azimuth. b) Coordinates describing the image: the migrated location $(\tilde{x}_m, \tilde{y}_m, z_m)$ or (x_m, y_m, z_m) with (for qP waves) $z_m = V_{P0} t_m/2$ and t_m the migrated two-way traveltime, the wavevectors $\boldsymbol{\xi}_s$ and $\boldsymbol{\xi}_r$ from the source (s) and receiver (r), and the wavevector associated with the reflector $\boldsymbol{\xi}_m$ (i.e. the dip covector).

2.3.1 The double-square-root equation

The double-square-root (DSR) equation, governing traveltime in a homogeneous isotropic medium, in 3D is given by

$$t_u = \frac{1}{v}\left(\sqrt{(\widetilde{x}_u - \widetilde{x}_m - h\sin\alpha)^2 + (\widetilde{y}_u - \widetilde{y}_m - h\cos\alpha)^2 + \left(\frac{vt_m}{2}\right)^2} + \sqrt{(\widetilde{x}_u - \widetilde{x}_m + h\sin\alpha)^2 + (\widetilde{y}_u - \widetilde{y}_m + h\cos\alpha)^2 + \left(\frac{vt_m}{2}\right)^2}\right), \tag{2.1}$$

where $\widetilde{x}_u$ and $\widetilde{y}_u$ are the common-midpoint (CMP) coordinates, $\widetilde{x}_m$ and $\widetilde{y}_m$ are the reflection-point coordinates, t_m is the two-way migrated traveltime, α is the acquisition azimuth measured positive in the direction of the positive $\widetilde{x}$ axis, v is the velocity, and h is the half-offset (see Figure 2.1). Note that here we consider only pure modes, i.e. P-P or S-S reflections, since the velocity at the source and receiver are assumed to be equal; the case of mode-converted waves should, however, be a straightforward generalization of the approach outlined in this section.

Rotating the positive $\widetilde{y}$ direction to the source-to-receiver direction, the DSR equation becomes

$$t_u = \frac{1}{v}\left(\sqrt{(x_u - x_m)^2 + (y_u - y_m - h)^2 + \left(\frac{vt_m}{2}\right)^2} + \sqrt{(x_u - x_m)^2 + (y_u - y_m + h)^2 + \left(\frac{vt_m}{2}\right)^2}\right). \tag{2.2}$$

To find x_u, y_u, t_u, p_u^x and p_u^y from $\widetilde{x}_u, \widetilde{y}_u, t_u, \widetilde{p}_u^x$ and $\widetilde{p}_u^y$, where p_u^x, p_u^y and $\widetilde{p}_u^x$, $\widetilde{p}_u^y$ are the horizontal slownesses of the unmigrated reflection in the rotated and unrotated coordinate systems, respectively, we use

$$\begin{pmatrix} x_u \\ y_u \\ t_u \\ p_u^x \\ p_u^y \end{pmatrix} = \begin{pmatrix} \cos\alpha & \sin\alpha & 0 & 0 & 0 \\ -\sin\alpha & \cos\alpha & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \cos\alpha & -\sin\alpha \\ 0 & 0 & 0 & \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} \widetilde{x}_u \\ \widetilde{y}_u \\ t_u \\ \widetilde{p}_u^x \\ \widetilde{p}_u^y \end{pmatrix}. \tag{2.3}$$

Table 2.1. Summary of notation.

variable	description
u, m	subscripts denoting migrated (m) and unmigrated (u) variables
s, r	subscripts denoting source (s) and receiver (r) variables
x, y, z	horizontal location x and y, and depth z
$t_{u,m}$	two-way unmigrated (u) and migrated (m) traveltime
$t_{s,r}$	one-way traveltime from source (s) or receiver (r) to reflection point
h	half-offset
v	group velocity
V	phase velocity
ψ	group angle
θ	phase angle
$p^{x,y}$	horizontal slowness in x or y direction
α	acquisition azimuth (measured anti-clockwise with positive y-axis)
$\phi_{x,y}$	reflector dip angle with horizontal in x or y direction
$\nu_{x,y}$	$-\tan\phi_{x,y}$
$\boldsymbol{\xi}_m$	dip covector (i.e., the wavevector associated with the reflector in the image)
$V_{P0,S0}$	vertical phase velocity for qP and qSV waves, respectively
$V_{NMO}(0)$	zero-dip NMO velocity
ϵ, δ	Thomsen parameters
η	anellipticity parameter
γ	azimuth angle of slowness vector with positive x-axis (clock-wise positive)
$(s,r)_\gamma$	$(\sin\gamma_s, \sin\gamma_r)$
$(s,r)_\theta$	$(\sin\theta_s, \sin\theta_r)$

To be consistent with the general treatment of time migration using common-midpoint coordinates and offset, we derive our results in this reference frame. Since we align the positive y-axis with the source-to-receiver direction, we develop our equations in the common offset, common azimuth domain. In the remaining text we assume the velocity to be known and equal to the RMS velocity; i.e. we develop the pre-stack map migration and demigration equations and their solutions in the context of time migration. Table 2.1 summarizes our notation throughout the remaining text.

2.3.2 Pre-stack common-offset migration

Equation (2.2) has three unknowns — x_m, y_m, and t_m. We obtain two additional equations by calculating the partial derivatives of t_u with respect to x_u and y_u while keeping the reflector location and offset constant, i.e., $p_u^x = \frac{1}{2}\frac{\partial t_u}{\partial x_u}$ and $p_u^y = \frac{1}{2}\frac{\partial t_u}{\partial y_u}$:

$$p_u^x = \frac{1}{2v}\left(\frac{x_u - x_m}{\sqrt{(x_u - x_m)^2 + (y_u - y_m - h)^2 + \left(\frac{vt_m}{2}\right)^2}} + \frac{x_u - x_m}{\sqrt{(x_u - x_m)^2 + (y_u - y_m + h)^2 + \left(\frac{vt_m}{2}\right)^2}}\right), \tag{2.4}$$

$$p_u^y = \frac{1}{2v}\left(\frac{y_u - y_m - h}{\sqrt{(x_u - x_m)^2 + (y_u - y_m - h)^2 + \left(\frac{vt_m}{2}\right)^2}} + \frac{y_u - y_m + h}{\sqrt{(x_u - x_m)^2 + (y_u - y_m + h)^2 + \left(\frac{vt_m}{2}\right)^2}}\right), \tag{2.5}$$

where the horizontal slownesses p_u^x and p_u^y can be measured. With these additional equations we arrive at a system of three equations with three unknowns. To derive equations (2.4) and (2.5) we used the property that on the pre-stack migration isochrone, defined by the DSR equation, $\frac{\partial t_m}{\partial x_u} = \frac{\partial t_m}{\partial y_u} = 0$ for constant h.

Solving equations (2.2), (2.4), and (2.5) for x_m, y_m and t_m results in

$$x_m = x_u - \frac{v^2 p_u^x t_u}{2}\left(1 - \Lambda_u^2\right), \tag{2.6}$$

$$y_m = y_u - \left(\frac{vt_u}{2}\right)^2 \frac{\Lambda_u}{h}, \tag{2.7}$$

$$t_m = 2\left\{\left(\frac{t_u}{2}\right)^2\left(1 - (vp_u^x)^2\right) - \left(\frac{h}{v}\right)^2 + \left(\frac{vt_u\Lambda_u}{4h}\right)^2\left(8\left(p_u^x h\right)^2 - t_u^2 + \left(\frac{2h}{v}\right)^2\left[1 - \left(vp_u^x\Lambda_u\right)^2\right]\right)\right\}^{\frac{1}{2}}, \tag{2.8}$$

in which

$$\Lambda_u = \Lambda_u\left(p_u^y, \Theta_u, h\right) \equiv \frac{1}{2\sqrt{2}p_u^y h}\sqrt{\Theta_u\left(1 - \sqrt{1 - \frac{64\left(p_u^y h\right)^4}{\Theta_u^2}}\right)}, \tag{2.9}$$

with

$$\Theta_u = \Theta_u\left(t_u, p_u^y, h\right) \equiv t_u^2 + \left(\frac{2h}{v}\right)^4 \frac{1}{t_u^2} - 2\left(\frac{2h}{v}\right)^2\left(1 - (vp_u^y)^2\right), \tag{2.10}$$

where the signs of the roots are chosen such that migration moves energy up-dip. The local dip of the reflector $p_m^{x,y} = \frac{1}{2}\frac{\partial t_m}{\partial (x,y)_m}$ can be found by calculating the partial derivatives $\left(\frac{\partial}{\partial x_m}, \frac{\partial}{\partial y_m}\right)$ of equation (2.2), using that $\frac{\partial t_u}{\partial x_m} = \frac{\partial t_u}{\partial y_m} = 0$ for constant h, and using equations (2.6)-(2.8) for x_m, y_m and t_m. This yields

$$\begin{aligned} p_m^{x,y} &= \frac{1}{2} p_u^{x,y} t_u \left|\Lambda_u - 1\right| \left|\Lambda_u + 1\right| \left\{ \left(\frac{t_u}{2}\right)^2 \left(1 - (v p_u^x)^2\right) - \left(\frac{h}{v}\right)^2 \right. \\ &\quad \left. + \left(\frac{v t_u \Lambda_u}{4h}\right)^2 \left(8\left(p_u^x h\right)^2 - t_u^2 + \left(\frac{2h}{v}\right)^2 \left(1 - (v p_u^x \Lambda_u)^2\right) \right) \right\}^{-\frac{1}{2}} . \end{aligned} \tag{2.11}$$

Equations (2.6)-(2.8) and (2.11) thus are explicit expressions that determine the migrated reflector coordinates $(x_m, y_m, t_m, p_m^x, p_m^y)$ from the specular reflection coordinates $(x_u, y_u, t_u, p_u^x, p_u^y)$, given h and v. Setting $x_m = x_u = 0$ and $p_m^x = p_u^x = 0$, reduces equations (2.7), (2.8) and (2.11) to the 2D equivalent expressions. Appendix C contains a similar analysis to the one presented here to derive the map dip moveout (DMO) equations.

Note that equations (2.6)-(2.8) and (2.11) do not use the offset horizontal slowness $p_h = \frac{1}{2}\frac{\partial t_u}{\partial h}$. This means that, in practice, only p_u^x and p_u^y need to be estimated, and the slope in a common-midpoint gather can be ignored. Usually expressions or algorithms for map migration use the slopes in both the common-offset and midpoint gathers, or, alternatively, the slopes in source and receiver gathers.

2.3.3 Zero-offset migration

Equations (2.6)-(2.8) and (2.11) seem to be singular at first sight for $h = 0$ and $p_u^h = 0$, and are strictly valid when $h \neq 0$ and $p_u^y \neq 0$. However, these singularities are introduced in the derivation of these expressions through division by h and p_u^y. For small h or p_u^y, i.e.,

$$\frac{64(p_u^y h)^4}{\Theta_u^2} \ll 1 ,$$

we use a first-order Taylor expansion for Λ_u such that

$$\Lambda_u \simeq \frac{2 p_u^y h}{\sqrt{\Theta_u}} . \tag{2.12}$$

Substituting this approximation for Λ_u into equations (2.6)-(2.8) and (2.11) gives

$$x_m \simeq x_u - \frac{v^2 p_u^x t_u}{2}\left(1 - \frac{4\left(p_u^y h\right)^2}{\Theta_u}\right), \tag{2.13}$$

$$y_m \simeq y_u - \frac{\left(v t_u\right)^2 p_u^y}{2\sqrt{\Theta_u}}, \tag{2.14}$$

$$\begin{aligned} t_m \simeq 2\Bigg\{ & \frac{t_u^2}{4}\left(1 - (v p_u^x)^2\right) - \left(\frac{h}{v}\right)^2 \\ & + \frac{1}{\Theta_u}\left(\frac{v t_u p_u^y}{2}\right)^2 \left(8\left(p_u^x h\right)^2 - t_u^2 + \left(\frac{2h}{v}\right)^2 \left[1 - \frac{1}{\Theta_u}\left(2 v p_u^x p_u^y h\right)^2\right]\right)\Bigg\}^{\frac{1}{2}}, \end{aligned} \tag{2.15}$$

$$\begin{aligned} p_m^{x,y} \simeq & \frac{1}{2} p_u^{x,y} t_u \\ & \times \left|2 p_u^y h/\sqrt{\Theta_u} - 1\right| \left|2 p_u^y h/\sqrt{\Theta_u} + 1\right| \Bigg\{ \left(\frac{t_u^2}{4}\left(1 - (v p_u^x)^2\right) - \left(\frac{h}{v}\right)^2\right) \\ & + \frac{1}{\Theta_u}\left(\frac{v t_u p_u^y}{2}\right)^2 \left(8\left(p_u^x h\right)^2 - t_u^2 + \left(\frac{2h}{v}\right)^2 \left[1 - \frac{1}{\Theta_u}\left(2 v p_u^x p_u^y h\right)^2\right]\right)\Bigg\}^{-\frac{1}{2}}. \end{aligned} \tag{2.16}$$

For the special case when $h = 0$, these equations reduce to their zero-offset (or post-stack) counterparts, i.e.,

$$(x, y)_m = (x, y)_u - \frac{v^2 p_u^{x,y} t_u}{2}, \tag{2.17}$$

$$t_m = t_u \sqrt{1 - v^2 p_u^2}, \tag{2.18}$$

$$p_m^{x,y} = \frac{p_u^{x,y}}{\sqrt{1 - v^2 p_u^2}}, \tag{2.19}$$

where we have defined

$$p_u \equiv \sqrt{(p_u^x)^2 + (p_u^y)^2}. \tag{2.20}$$

Equations (2.17)-(2.18) can also be found in Haas & Viallix (1976); note the typographical error in their equations for the migrated location. Setting $x_m = x_u = 0$ and $p_m^x = p_u^x = 0$ gives the expressions in 2D. These 2D expressions can also be found in Claerbout (1985, Chapter 1).

2.3.4 Pre-stack common-offset demigration

The demigration equations for x_u and y_u can be found by first solving equation (2.2) for t_m and evaluating the partial derivatives $\frac{\partial t_m}{\partial x_m}$ and $\frac{\partial t_m}{\partial y_m}$. Using the resulting expressions

for x_u and y_u in equation (2.2) then give the explicit expression for t_u. To find the slopes p_u^x and p_u^y, we then simply substitute the expressions for x_u, y_u, and t_u in equations (2.4) and (2.5). The resulting equations are

$$x_u = x_m + \frac{v^2 p_m^x t_m}{2}, \tag{2.21}$$

$$y_u = y_m + \frac{v^2 p_m^y t_m}{2} + h\Lambda_m, \tag{2.22}$$

$$t_u = \sqrt{\frac{4h^2}{v^2} + \frac{2p_m^y h t_m}{\Lambda_m}}, \tag{2.23}$$

$$p_u^{x,y} = \frac{p_m^{x,y} t_m}{|\Lambda_m - 1|\,|\Lambda_m + 1| \sqrt{\frac{4h^2}{v^2} + \frac{2p_m^y h t_m}{\Lambda_m}}}, \tag{2.24}$$

in which

$$\Lambda_m = \Lambda_m\left(p_m^y, \Theta_m, h\right) \equiv \frac{4p_m^y h}{\Theta_m\left(1 + \sqrt{1 + \frac{16(p_m^y h)^2}{\Theta_m^2}}\right)}, \tag{2.25}$$

with

$$\Theta_m = \Theta_m(t_m, p_m^{x,y}) \equiv t_m\left(1 + v^2\left[(p_m^x)^2 + (p_m^y)^2\right]\right). \tag{2.26}$$

Equations (2.21)-(2.24) determine the specular reflection $(x_u, y_u, t_u, p_u^x, p_u^y)$ from the reflector $(x_m, y_m, t_m, p_m^x, p_m^y)$. Note that $p_m^{x,y}$ can be estimated from the dip of the imaged reflector using that

$$\tan\phi_{x,y} = \frac{v}{2}\frac{\partial t_m}{\partial\,(x,y)_m} = v p_m^{x,y}, \tag{2.27}$$

where $\phi_{x,y}$ is the reflector dip angle with the horizontal in the x- or y-direction (measured positive clockwise). Again, the 2D case follows by setting $x_m = x_u = 0$ and $p_m^x = p_u^x = 0$. The inverse map DMO equations are derived in Appendix D through a similar analysis to the one presented here.

2.3.5 Zero-offset demigration

The demigration mapping given by equations (2.21)-(2.24) indeed reduces to its zero-offset counterpart if $h = 0$. The resulting expressions are

$$(x,y)_u = (x,y)_m + \frac{v^2 p_m^{x,y} t_m}{2}, \tag{2.28}$$

$$t_u = t_m\sqrt{1 + v^2 p_m^2}, \tag{2.29}$$

$$p_u^{x,y} = \frac{p_m^{x,y}}{\sqrt{1 + v^2 p_m^2}}, \tag{2.30}$$

where

$$p_m \equiv \sqrt{(p_m^x)^2 + (p_m^y)^2}\ . \tag{2.31}$$

With $x_m = x_u = 0$ and $p_m^x = p_u^x = 0$ these expressions reduce to the 2D case.

2.3.6 Numerical example

Figures 2.2a and b show modeled common-offset reflection seismic data from a syncline reflector, embedded in a constant velocity medium with $v = 2000$ m/s, and its migrated counterpart, respectively; the offset is 2000 m (i.e., $h = 1000$ m). To verify the pre-stack map migration and demigraton equations numerically, Figures 2.2c and d show line elements overlaid on top of the common-offset data and the migrated data, respectively. The line elements in Figure 2.2c were obtained using points along the syncline in our model (giving y_m and t_m), and the associated local dips at these points. These dips were converted to migrated horizontal slowness (i.e., p_m^y) using equation (2.27), and the values of y_u, t_u, and p_u^y, i.e., the demigrated variables, were found using the 2D equivalent expressions of equations (2.22)-(2.24) (i.e., with $p_m^x = 0$). To find the line elements associated with the diffractions from the edges of the syncline, we took a fan of line elements centered at these edges, with a range of dips from -80 to 80 degrees, and demigrated them according to the procedure just described. Figure 2.2c shows excellent agreement between the line elements and the data, indicating the validity of the common-offset map time-demigration equations (2.21)-(2.24).

The values of y_u, t_u, and p_u^y thus obtained, were subsequently used in equations (2.7), (2.8), and (2.11) (with $p_u^x = 0$) to calculate the migrated counterparts y_m, t_m, and p_m^y. The obtained values for p_m^y were converted to local reflector dip using equation (2.27), giving the orientations of the line elements shown in Figure 2.2d. The resulting line elements coincide with the migrated data, indicating the validity of the common-offset map time-migration equations (2.6)-(2.8) and (2.11). Note how the line elements from the diffractions (indicated by the black line elements) associated with the edges of the syncline shaped reflector, are all map migrated to the same location but with different orientations; much like the Fourier transform of a delta function in space has all $\boldsymbol{k}$-directions.

2.4 Map time-migration and demigration in VTI media

We consider now the case of homogeneous VTI media. In our derivation, we make use of the fact that in such media the phase and group velocity vectors lie in the vertical plane. Since for homogeneous TTI media this requirement is also satisfied in the vertical symmetry plane that contains the symmetry axis, our results also apply to this plane in homogeneous TTI media. Furthermore, the kinematic equivalence of TI media and orthorhombic media in the symmetry planes, generalizes our results to these planes in orthorhombic media.

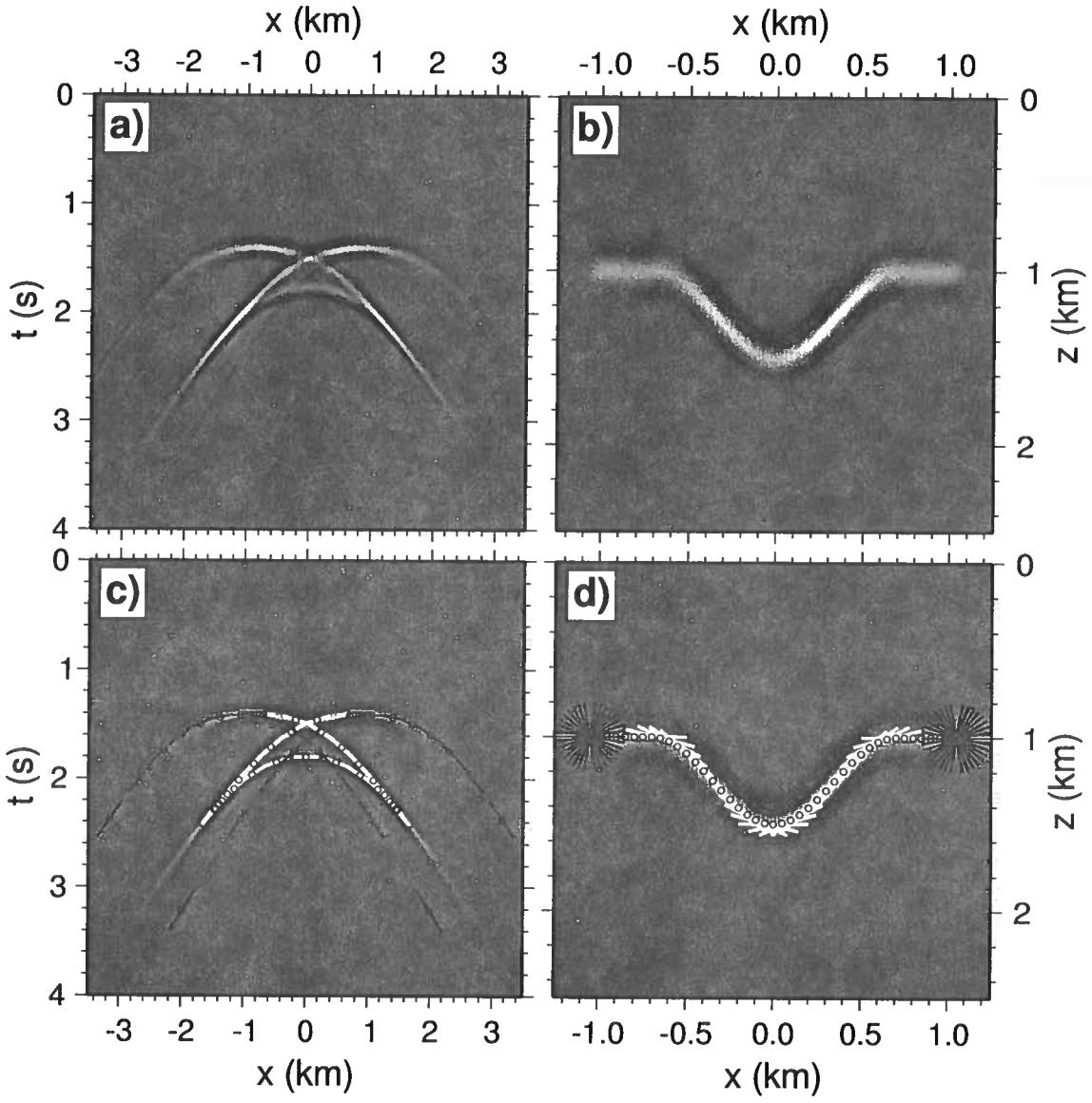


Figure 2.2. Common offset ($h = 1000$ m) data (a) and its image (b) from a syncline shaped reflector embedded in a constant velocity ($v = 2000$ m/s) medium, and demigrated (c) and migrated (d) line elements superposed on the data and image, respectively. The excellent agreement between the demigrated line elements and the seismic data (c), and the migrated line elements and the image (d), indicate the validity of the common-offset map time-demigration and migration equations, respectively.

2.4.1 The double-square-root equation

In general, the group velocity vector is perpendicular to the slowness surface, whereas the slowness vector is perpendicular to the wavefront. In TI media, the group velocity v depends only on the phase angle θ with the axis of rotational symmetry, and is given by (Tsvankin, 2001, p.29)

$$v = V(\theta)\sqrt{1+\left(\frac{1}{V(\theta)}\frac{dV}{d\theta}\right)^2}\,, \tag{2.32}$$

where V is the phase velocity of either qP or qSV waves. The group angle ψ in such media follows from

$$\tan\psi = \frac{\tan\theta + \dfrac{1}{V(\theta)}\dfrac{dV}{d\theta}}{1-\dfrac{\tan\theta}{V(\theta)}\dfrac{dV}{d\theta}}\,. \tag{2.33}$$

The group angle ψ is defined as the angle of the ray with the rotational symmetry axis, while the phase angle θ is the angle of the normal to the wavefront with the symmetry axis (see Figure 2.1a). Using that the energy travels along a ray with the group velocity, the DSR equation for a homogeneous anisotropic medium is given by

$$t_u = \frac{\sqrt{(x_{s,r}-x_m)^2+(y_s-y_m)^2+z_m^2}}{v_s} + \frac{\sqrt{(x_{s,r}-x_m)^2+(y_r-y_m)^2+z_m^2}}{v_r}\,, \tag{2.34}$$

where $v_{s,r}$ are the group velocities in the directions of the rays connecting the source with the reflection point, and the receiver with the reflection point, respectively, and $(x_{s,r}, y_{s,r})$ are the source and receiver coordinates. The positive y-direction is in the source-to-receiver direction (i.e., $x_{s,r} = x_u$), and the coordinate system is right-handed as before (see Figure 2.1a). Note that $v_{s,r}$ could be either the group velocity for qP or qSV waves.

2.4.2 Medium parametrization

For general transversely isotropic (TI) media, the phase velocity for qP and qSV waves using the parameterization introduced by Thomsen (1986), is given by (Tsvankin, 2001, p.22)

$$V(\theta) = V_{P0}\sqrt{1+\epsilon\sin^2\theta - \frac{f}{2} \pm \frac{f}{2}\sqrt{\left(1+\frac{2\epsilon\sin^2\theta}{f}\right)^2 - \frac{2(\epsilon-\delta)\sin^2 2\theta}{f}}}\,, \tag{2.35}$$

where V_{P0} is the phase velocity for the qP wave at $\theta = 0$, $f \equiv 1 - \frac{V_{S0}^2}{V_{P0}^2}$ with V_{S0} the phase velocity of the qSV wave at $\theta = 0$, and ϵ and δ are the Thomsen anisotropy parameters. The plus sign refers to qP waves, and the minus sign to qSV waves. Taking the derivative

with respect to the phase angle θ gives

$$\frac{dV}{d\theta} = \frac{V_{P0}^2}{V(\theta)} \left(\epsilon s\sqrt{1-s^2} \pm \frac{\left[\left(1+\frac{2\epsilon s^2}{f}\right)\epsilon s\sqrt{1-s^2} - 2\left(\epsilon-\delta\right)s\sqrt{1-5s^2-4s^6}\right]}{\sqrt{\left(1+\frac{2\epsilon s^2}{f}\right)^2 - \frac{8\left(\epsilon-\delta\right)s^2\left(1-s^2\right)}{f}}} \right), \tag{2.36}$$

where $s \equiv \sin\theta$, and again the plus sign refers to qP waves and the minus sign to qSV waves. Note that we have $0 \leq \theta < \pi/2$, since rays cannot turn in homogeneous media.

The qP-wave phase velocity, however, depends only weakly on the vertical shear-wave velocity V_{S0} [e.g., Tsvankin (1996) and Alkhalifah (1998); a precise analysis of this is contained in Schoenberg & de Hoop (2000)], such that the influence of V_{S0} on all kinematic problems involving qP waves can be ignored. Because in this paper we are only dealing with the geometry of map migration, we can for most practical purposes set $f = 1$ in equations (2.35), (2.36), and (2.53). If V_{S0} is known, f can be calculated and subsequently used in equations (2.35), (2.36) and (2.53) to find the phase velocity, its derivative, and the phase angles.

Alkhalifah & Tsvankin (1995) showed that the time signatures (e.g., reflection moveout, DMO, and time-migration operators) of qP waves in homogeneous VTI media are mainly characterized by the zero-dip normal-moveout (NMO) velocity $V_{NMO}(0) = V_{P0}\sqrt{1+2\delta}$ and the anellipticity parameter $\eta = (\epsilon-\delta)/(1+2\delta)$, with an almost negligible influence of V_{P0}.[2] Using these expressions for η and $V_{NMO}(0)$, equations (2.35) and (2.36) can, for qP waves, be rewritten in terms of η, $V_{NMO}(0)$, and V_{P0}. The expressions we derive for map migration in VTI media can be used to estimate the anellipticity parameter η and $V_{NMO}(0)$, by using the slope information of one event at two (or more) different offsets, and calculating the migrated times for these offsets for assumed values of η and $V_{NMO}(0)$ (given some V_{P0}). The correct values of η and $V_{NMO}(0)$ should yield the same migrated time for all offsets, since the data have one common reflection point.

2.4.3 Pre-stack map time-migration

For VTI media, all vertical planes are medium mirror symmetry planes. Both vertical planes, the one defined by the source position and the reflector position, and the one defined by the receiver position and the reflector position, thus are symmetry planes also. Throughout the remainder, we refer to these planes as the *source* and *receiver planes* (see Figure 2.1). In the source plane,

$$\tan\psi_s = \frac{\sqrt{(x_s-x_m)^2+(y_s-y_m)^2}}{z_m}, \tag{2.37}$$

[2]To avoid confusion with the NMO velocity at finite dip, we prefer to maintain the notation $V_{NMO}(0)$ rather than writing V_{NMO} as some authors do.

while in the receiver plane, we have

$$\tan\psi_r = \frac{\sqrt{(x_r - x_m)^2 + (y_r - y_m)^2}}{z_m}\,, \tag{2.38}$$

where $\psi_{s,r}$ are the group angles at the source and receiver, and z_m is the migrated depth. The horizontal slownesses satisfy the relation

$$p_{s,r} = \frac{\sin\theta_{s,r}}{V_{s,r}}\,, \tag{2.39}$$

where we defined

$$V_{s,r} \equiv V(\theta_{s,r})\,, \tag{2.40}$$

$$p_{s,r} \equiv \sqrt{\left(p^x_{s,r}\right)^2 + \left(p^y_{s,r}\right)^2}\,, \tag{2.41}$$

with $p^{x,y}_{s,r}$ denoting the horizontal slownesses at the source or receiver in the x or y direction, and $\theta_{s,r}$ the phase angle at the source or receiver. Then, using equation (2.39) in equation (2.33) and substituting the result in equations (2.37) and (2.38), we get

$$\frac{\sqrt{(x_{s,r} - x_m)^2 + (y_{s,r} - y_m)^2}}{z_m} = \frac{\dfrac{V_{s,r}p_{s,r}}{\sqrt{1 - V^2_{s,r}p^2_{s,r}}} + \dfrac{1}{V_{s,r}}\left.\dfrac{dV}{d\theta}\right|_{s,r}}{1 - \dfrac{p_{s,r}}{\sqrt{1 - V^2_{s,r}p^2_{s,r}}}\left.\dfrac{dV}{d\theta}\right|_{s,r}}\,. \tag{2.42}$$

Here $\left.\frac{dV}{d\theta}\right|_{s,r}$ is the derivative of the phase velocity with respect to the phase angle θ with the vertical symmetry axis, evaluated at the phase angle at the source (θ_s) or receiver (θ_r).

Using equation (2.42) in (2.34) then results in an expression for the migrated depth

$$z_m = t_u \left(\frac{1}{V_s\left(\sqrt{1 - V_s^2p_s^2} - p_s\left.\frac{dV}{d\theta}\right|_s\right)} + \frac{1}{V_r\left(\sqrt{1 - V_r^2p_r^2} - p_r\left.\frac{dV}{d\theta}\right|_r\right)} \right)^{-1}\,. \tag{2.43}$$

For pure mode waves, i.e., qP-qP or qSV-qSV, the migrated depth can be converted to two-way migrated time $t_m = 2z_m/V_{P0,S0}$. Defining $\gamma_{s,r}$ as the angles of the horizontal projection of the slowness vector at the source and receiver with the positive x-axis (see Figure 2.3), we find

$$\sin\gamma_{s,r} = \frac{y_{s,r} - y_m}{\sqrt{(x_{s,r} - x_m)^2 + (y_{s,r} - y_m)^2}} = \frac{p^y_{s,r}}{p_{s,r}}\,. \tag{2.44}$$

Using equation (2.43) in (2.42) to get an expression for $\sqrt{(x_{s,r} - x_m)^2 + (y_{s,r} - y_m)^2}$, and

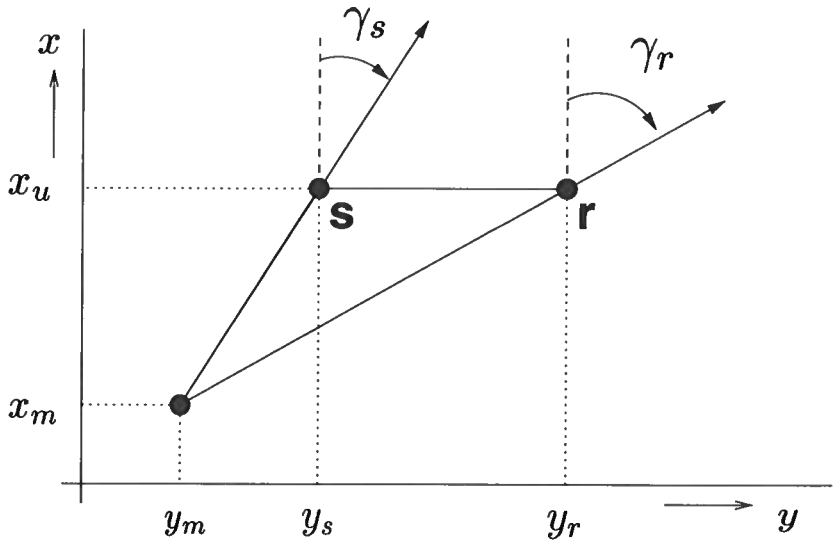


Figure 2.3. Definition of $\gamma_{s,r}$ as the angles of the horizontal projections of the slowness vectors with the positive x-axis.

substituting the resulting expression in equation (2.44), then gives

$$y_m = y_{s,r} - \frac{t_u p^y_{s,r} \left(V_{s,r} + \sqrt{\frac{1}{V^2_{s,r} p^2_{s,r}} - 1}\, \frac{dV}{d\theta}\Big|_{s,r} \right) \left(\sqrt{1 - V^2_{r,s} p^2_{r,s}} - p_{r,s}\, \frac{dV}{d\theta}\Big|_{r,s} \right)}{\frac{1}{V_r} \left(\sqrt{1 - V^2_s p^2_s} - p_s\, \frac{dV}{d\theta}\Big|_s \right) + \frac{1}{V_s} \left(\sqrt{1 - V^2_r p^2_r} - p_r\, \frac{dV}{d\theta}\Big|_r \right)} . \quad (2.45)$$

Note that either the parameters at the source or receiver can be used to calculate y_m, but that using the source parameters in the first term of the numerator implies using the receiver parameters in the second term, and vice versa; hence the order of the subscripts s, r and r, s. To find x_m, we first calculate

$$\cos \gamma_{s,r} = \frac{x_{s,r} - x_m}{\sqrt{(x_{s,r} - x_m)^2 + (y_{s,r} - y_m)^2}} = \frac{p^x_{s,r}}{p_{s,r}} . \quad (2.46)$$

Then, using equation (2.43) again in (2.42) to get an expression for

$\sqrt{(x_{s,r}-x_m)^2+(y_{s,r}-y_m)^2}$, and using the result in (2.46), gives

$$x_m = x_{s,r} - \frac{t_u p^x_{s,r}\left(V_{s,r}+\sqrt{\frac{1}{V^2_{s,r}p^2_{s,r}}-1}\left.\frac{dV}{d\theta}\right|_{s,r}\right)\left(\sqrt{1-V^2_{r,s}p^2_{r,s}}-p_{r,s}\left.\frac{dV}{d\theta}\right|_{r,s}\right)}{\frac{1}{V_r}\left(\sqrt{1-V_s^2p_s^2}-p_s\left.\frac{dV}{d\theta}\right|_s\right)+\frac{1}{V_s}\left(\sqrt{1-V_r^2p_r^2}-p_r\left.\frac{dV}{d\theta}\right|_r\right)} . \quad (2.47)$$

Again note the order of the subscripts s,r and r,s. Thus, equations (2.43), (2.45), and (2.47) are closed-form expressions for the migrated location. The phase velocity and its derivative at the source and receiver in equations (2.43), (2.45) and (2.47) can be set to their respective values for either qP or qSV waves, showing the applicability of these expressions for both pure modes and mode-converted waves. The 2D expressions follow by setting $x_m = x_u = 0$ and $p^x_s = p^x_r = 0$.

To find the reflector dip covector $\boldsymbol{\xi}_m$ (i.e., the wave-vector associated with the reflector in the image, see Figure 2.1b), we use that the slowness vectors $\boldsymbol{p}_{s,r}$ obey Snell's law upon reflection at the reflector. Since we define the vertical components of the slowness vectors to point in the negative z-direction (i.e., upwards), we have

$$\boldsymbol{\xi}_m = \boldsymbol{\xi}_s + \boldsymbol{\xi}_r = -\omega\left(\boldsymbol{p}_s + \boldsymbol{p}_r\right) , \quad (2.48)$$

where ω is the angular frequency, and $\boldsymbol{\xi}_{s,r}$ are the wave vectors associated with the source and receiver rays. The slowness vectors at the source and receiver are given by

$$\boldsymbol{p}_{s,r} = \begin{pmatrix} p^x_{s,r} \\ p^y_{s,r} \\ -\sqrt{\frac{1}{V^2_{s,r}} - p^2_{s,r}} \end{pmatrix} . \quad (2.49)$$

Therefore, the dip covector $\boldsymbol{\xi}_m$ is given by

$$\boldsymbol{\xi}_m = \omega \begin{pmatrix} -\left(p^x_s + p^x_r\right) \\ -\left(p^y_s + p^y_r\right) \\ \sqrt{\frac{1}{V_s^2} - p_s^2} + \sqrt{\frac{1}{V_r^2} - p_r^2} \end{pmatrix} . \quad (2.50)$$

This expressions holds for pure modes and mode-converted waves.

For pure modes, $\boldsymbol{\xi}_m$ can be translated to the migrated horizontal slowness components $p^{x,y}_m$, using that (defining the x and y components $\nu_{x,y}$ of the dip)

$$-\nu_{x,y} \equiv \frac{-\xi^{x,y}_m}{\xi^z_m} = \tan\phi_{x,y} = \frac{V_{P0,S0}}{2}\frac{\partial t_m}{\partial\left(x,y\right)_m} = V_{P0,S0}p^{x,y}_m , \quad (2.51)$$

where $\xi^{x,y,z}_m$ are the components of $\boldsymbol{\xi}_m$, and $\phi_{x,y}$ is again the reflector dip with the horizontal in the x- or y-direction (measured positive clockwise). Using equation (2.50) in (2.51), it

follows that

$$p_m^{x,y} = \frac{p_s^{x,y} + p_r^{x,y}}{V_{P0,S0}\left(\sqrt{\frac{1}{V_s^2} - p_s^2} + \sqrt{\frac{1}{V_r^2} - p_r^2}\right)} \,. \tag{2.52}$$

Of course, in equations (2.51) and (2.52) V_{P0} and V_{S0} are used for qP-qP and qSV-qSV waves respectively. Again, setting $x_m = x_u = 0$ and $p_s^x = p_r^x = 0$ gives the 2D expressions.

2.4.4 Source and receiver phase angles

Equations (2.35) and (2.36) are used to calculate the phase velocity and its derivative, that in turn are used in equations (2.43), (2.45), (2.47) and (2.50), to find the location and orientation of the reflector in the image. To calculate the phase velocity and its derivative using equations (2.35) and (2.36), the angle θ, and thus $s = \sin\theta$, needs to be known. To find s, we need to solve equation (2.39) for s using equation (2.35) for the phase velocity. This gives

$$\begin{aligned} \sin\theta_{s,r} \;=\; & \Bigg\{ P_{s,r}\left[(2-f) - 2P_{s,r}\left(\epsilon - \delta f\right)\right] \\ & \times \frac{\left[1 \pm \sqrt{1 - \frac{4\left(1-f\right)\left(1 - 2\epsilon P_{s,r} - 2P_{s,r}^2 f\left(\epsilon-\delta\right)\right)}{\left(f - 2 + 2P_{s,r}\left(\epsilon - \delta f\right)\right)^2}}\right]}{2\left(1 - 2\epsilon P_{s,r} - 2P_{s,r}^2 f\left(\epsilon - \delta\right)\right)} \Bigg\}^{\frac{1}{2}} \,, \end{aligned} \tag{2.53}$$

where

$$P_{s,r} \equiv p_{s,r}^2 V_{P0}^2 \,, \tag{2.54}$$

with $p_{s,r}$ defined in equation (2.41), the plus sign refering to qP waves, and the minus sign to qSV waves. Note that for VTI media in 2D we have $-\pi/2 < \theta_{s,r} < \pi/2$ and that the sign of $\sin\theta_{s,r}$ is given by the sign of $p_{s,r}$; in this case the right hand side of equation (2.53) is preceded by $\mathrm{sgn}(p_{s,r})$.

Therefore, given $V_{P0,S0}$ and the anisotropic parameters ϵ and δ [or η and $V_{NMO}(0)$ for qP waves], we can use equation (2.53) with the measured $p_{s,r}$ to calculate the phase angles $\theta_{s,r}$ for both qP and qSV waves. The resulting values can then be used in equations (2.35) and (2.36) to find the phase velocity and its derivative at both the source and receiver location. Note that for the vertical symmetry plane of TTI media we have $p = \sin(\nu + \theta)/V(\theta)$, with ν the angle with the vertical of the symmetry axis. In this case we solve numerically for θ provided ν is known.

If the horizontal slowness is used to parametrize the phase velocity and its derivative,

we need not solve for the phase angles. To find the phase velocity as a function of the horizontal slowness, we simply replace $\sin\theta_{s,r}$ with $V_{s,r}p_{s,r}$ in equation (2.35) and solve for $V_{s,r}$. In Appendix A, we give the resulting expressions for the phase velocity and its derivative as functions of the horizontal slowness for qP waves, using that $f = 1$ for most practical purposes.

2.4.5 Zero-offset migration

By setting $p_s^{x,y} = p_r^{x,y} = p_u^{x,y}$ and $\theta_s = \theta_r = \theta$, the pre-stack map migration equations reduce to their zero-offset counterparts. Here we treat pure modes only, i.e., we set $V_s = V_r = V(\theta)$. Doing this for equations (2.43), (2.45), (2.47), (2.50), and (2.52) gives

$$z_m = \frac{V(\theta)t_u}{2}\left(\sqrt{1-V^2(\theta)p_u^2} - p_u \left.\frac{dV}{d\theta}\right|_\theta\right), \tag{2.55}$$

$$(x,y)_m = (x,y)_u - \frac{V^2(\theta)p_u^{x,y}t_u}{2} - \frac{V(\theta)p_u^{x,y}t_u}{2}\sqrt{\frac{1}{V^2(\theta)p_u^2}-1}\left.\frac{dV}{d\theta}\right|_\theta, \tag{2.56}$$

$$\boldsymbol{\xi}_m = 2\omega \begin{pmatrix} -p_u^x \\ -p_u^y \\ \sqrt{\dfrac{1}{V^2(\theta)} - p_u^2} \end{pmatrix}, \tag{2.57}$$

$$p_m^{x,y} = \frac{V(\theta)p_u^{x,y}}{V_{P0,S0}\sqrt{1-V^2(\theta)p_u^2}}, \tag{2.58}$$

where p_u is defined in equation (2.20). Of course, the migrated depth z_m can be converted to two-way migrated traveltime $t_m = 2z_m/V_{P0,S0}$. Setting $x_m = x_u = 0$ and $p_m^x = p_u^x = 0$, the 2D expressions follow from their 3D counterparts. It can be shown that the 2D counterparts of equations (2.55) and (2.56) for the migrated time and location, are equivalent to equation (28) in Cohen (1998). Also, setting $\dfrac{dV}{d\theta} = 0$ and replacing $V(\theta)$ and V_{P0} with v, these expressions for VTI media reduce to their counterparts for isotropic media [cf. equations (2.17) - (2.19)].

2.4.6 Pre-stack map time-demigration

For the demigration problem, we assume the migrated location (x_m, y_m, z_m) and migrated dips $\phi_{x,y}$ (or $\nu_{x,y}$) are given. To find the unmigrated midpoint location, we need to find the phase angles with the (vertical) symmetry axis and the azimuths of the rays from both the source and the receiver to the reflection point. If we know the angles with the symmetry axis, we can use equations (2.35), (2.36), and (2.39) to find the phase velocity, its derivative and $p_{s,r}$. The projections $p_{s,r}^{x,y}$ are then calculated using

$$p_{s,r}^x = -\operatorname{sgn}(\nu_x)p_{s,r}\sqrt{1-(s,r)_\gamma^2}, \tag{2.59}$$

$$p_{s,r}^y = p_{s,r}\,(s,r)_\gamma, \tag{2.60}$$

with $(s,r)_\gamma = \sin\gamma_{s,r}$, the azimuth angles $\gamma_{s,r}$ defined in Figure 2.3, ν_x defined in equation (2.51), and $\mathrm{sgn}(\nu_x)$ denoting the sign of ν_x. The unmigrated location then follows from solving equations (2.43), (2.45), and (2.47) for t_u, y_u, and x_u, using the values for the phase velocity, its derivative, and $p^{x,y}_{s,r}$.

To find the azimuth angles, $\gamma_{s,r}$, and the angles with the vertical symmetry axis, $\theta_{s,r}$, we use the offset and azimuth information, and the dips $\phi_{x,y}$ or $\nu_{x,y} = -\tan\phi_{x,y}$ [see equation (2.51)]. Since in the rotated coordinate system (where the positive y-axis is in the source-to-receiver direction) the projection onto the x-axis of the ray connecting the source and the reflector equals the projection of the ray connecting the receiver and the reflector onto this axis, we must have

$$\sqrt{1-s_\gamma^2}\frac{\left(s_\theta+\dfrac{\sqrt{1-s_\theta^2}}{V_s}\left.\dfrac{dV}{d\theta}\right|_s\right)}{\sqrt{1-s_\theta^2}-\dfrac{s_\theta}{V_s}\left.\dfrac{dV}{d\theta}\right|_s}=\sqrt{1-r_\gamma^2}\frac{\left(r_\theta+\dfrac{\sqrt{1-r_\theta^2}}{V_r}\left.\dfrac{dV}{d\theta}\right|_r\right)}{\sqrt{1-r_\theta^2}-\dfrac{r_\theta}{V_r}\left.\dfrac{dV}{d\theta}\right|_r}\,, \tag{2.61}$$

where $(s,r)_\theta \equiv \sin\theta_{s,r}$ with $\theta_{s,r}$ the phase angles.

Note that since $0 \le \theta_{s,r} < \pi/2$ and $0 \le \gamma_{s,r} < 2\pi$, we have $0 \le (s,r)_\theta < 1$ and $-1 \le (s,r)_\gamma \le 1$. Because the azimuth angles vary between 0 and 2π, we need to keep track of the sign of $\cos\gamma_{s,r}$. In our rotated coordinate system, the sign of ν_x determines the sign of $\cos\gamma_{s,r}$. Therefore, for demigration,

$$\cos\gamma_{s,r} = -\,\mathrm{sgn}(\nu_x)\sqrt{1-(s,r)_\gamma^2}\,. \tag{2.62}$$

This also explains the presence of $\mathrm{sgn}(\nu_x)$ in equation (2.59).

Furthermore, the difference of the projections onto the y-axis should equal the offset $2h$, i.e.,

$$z_m\left(r_\gamma\frac{\left(r_\theta+\dfrac{\sqrt{1-r_\theta^2}}{V_r}\left.\dfrac{dV}{d\theta}\right|_r\right)}{\sqrt{1-r_\theta^2}-\dfrac{r_\theta}{V_r}\left.\dfrac{dV}{d\theta}\right|_r}-s_\gamma\frac{\left(s_\theta+\dfrac{\sqrt{1-s_\theta^2}}{V_s}\left.\dfrac{dV}{d\theta}\right|_s\right)}{\sqrt{1-s_\theta^2}-\dfrac{s_\theta}{V_s}\left.\dfrac{dV}{d\theta}\right|_s}\right)=2h\,. \tag{2.63}$$

Using the definitions of $p^{x,y}_{s,r}$ and equation (2.50) for $\boldsymbol{\xi}_m$ in equation (2.51) for $\nu_{x,y}$, it follows that

$$\nu_x = \mathrm{sgn}(\nu_x)\left(\frac{V_r s_\theta\sqrt{1-s_\gamma^2}+V_s r_\theta\sqrt{1-r_\gamma^2}}{V_r\sqrt{1-s_\theta^2}+V_s\sqrt{1-r_\theta^2}}\right)\,, \tag{2.64}$$

$$\nu_y = -\left(\frac{V_r s_\theta s_\gamma + V_s r_\theta r_\gamma}{V_r\sqrt{1-s_\theta^2} + V_s\sqrt{1-r_\theta^2}}\right) . \tag{2.65}$$

Equation (2.61) together with equations (2.63)-(2.65) form a system of four nonlinear equations with four unknowns: $(s,r)_\theta$ and $(s,r)_\gamma$. The phase velocity and its derivative at the source and receiver can be set to their respective values for either qP or qSV waves. Hence, these expressions are valid for both pure modes and mode-converted waves. Attempts to eliminate, for example, r_θ, s_γ, and r_γ to get one equation in s_θ lead to a high order polynomial equation in s_θ. Therefore, to find the unknown angles, a numerical scheme such as Gauss-Newton (Dennis, 1977) or Levenberg-Marquardt (Moré, 1977) might be used, with the isotropic solution as initial value. If $(s,r)_\theta$ and $(s,r)_\gamma$ are known, the scattering angle can be found using the cosine formula.

In Appendix A, we show that for qP waves the system of equation (2.61) and equations (2.63)-(2.65) can be rewritten into a somewhat simpler system using the horizontal slownesses instead of the phase angles, under the assumption that for most practical purposes we can set $f = 1$. Note that for the 2D problem, $(s,r)_\gamma = 1$ and $-1 < (s,r)_\theta < 1$, so equations (2.63) and (2.65) form a set of two nonlinear equations with two unknowns, $(s,r)_\theta$.

2.4.7 Zero-offset demigration

For zero-offset reflections in homogeneous anisotropic media, the angle of the phase direction with the vertical equals the reflector dip angle; for VTI media the phase angle thus equals this dip angle. Since in the demigration problem this dip is known, we know the phase angle and thus the group angle. Note that in the vertical symmetry plane of TTI media the reflector dip equals the sum of the phase angle and the tilt of the symmetry axis; for zero-offset demigration in such media the phase angle is thus also known, provided the tilt of the symmetry axis is known. Therefore, once the position and orientation of the reflector are known, we can calculate the unmigrated location. Here we treat pure modes only.

Using equation (2.58), we find

$$p_u^2 = \frac{V_{P0,S0}^2 p_m^2}{V^2(\theta)\left(1 + V_{P0,S0}^2 p_m^2\right)} , \tag{2.66}$$

with p_m defined in equation (2.31). Then, using this expression in equation (2.58), we get

$$p_u^{x,y} = \frac{V_{P0,S0} p_m^{x,y}}{V(\theta)\sqrt{1 + V_{P0,S0}^2 p_m^2}} . \tag{2.67}$$

To find the two-way traveltime t_u, we use equation (2.66) in (2.55) and solve for t_u. This

gives

$$t_u = \frac{2z_m\sqrt{1+V_{P0,S0}^2 p_m^2}}{\left(V(\theta) - p_m V_{P0,S0}\left.\frac{dV}{d\theta}\right|_\theta\right)}. \tag{2.68}$$

Then, using equations (2.66)-(2.68) in equation (2.56), gives the unmigrated location,

$$(x,y)_u = (x,y)_m + \frac{V_{P0,S0}p_m^{x,y} z_m\left(V(\theta) + \frac{1}{V_{P0,S0}p_m}\left.\frac{dV}{d\theta}\right|_\theta\right)}{\left(V(\theta) - V_{P0,S0}p_m\left.\frac{dV}{d\theta}\right|_\theta\right)}. \tag{2.69}$$

Note that by setting $\frac{dV}{d\theta} = 0$ and replacing $V(\theta)$ and $V_{P0,S0}$ with the constant velocity v, the equations for zero-offset demigration in VTI media reduce to their isotropic equivalents [cf. equations (2.28)-(2.30)]. Also, setting $x_m = x_u = 0$ and $p_m^x = p_u^x = 0$ gives the expressions in 2D.

Since for zero-offset data in VTI homogeneous media the reflector dip equals the phase angle, we find the phase angle θ from

$$\theta = \arctan\sqrt{\nu_x^2 + \nu_y^2}. \tag{2.70}$$

The calculated value of θ can subsequently be used in equation (2.35) and (2.36) to find the phase velocity and its derivative. Again, we can either use $V(\theta)$ and $\left.\frac{dV}{d\theta}\right|_\theta$ for qP or qSV waves. For qP waves we use V_{P0} in equations (2.66)-(2.69), while for qSV waves we use V_{S0}.

2.5 The applicability of map depth-migration and demigration in heterogeneous anisotropic media

The closed-form expressions we derived for common-offset and common-azimuth map time-migration in homogeneous isotropic, and for common-azimuth map time-migration in VTI media, explicitly show that, for a given offset and azimuth, the mapping from the surface measurements (i.e., source and receiver positions, traveltimes and slopes), to the subsurface image (i.e., reflector position and orientation), is *one-to-one* for such media (assuming the recorded wavefields reflected only once in the subsurface). Given a scattering angle and azimuth, this mapping exists also in heterogeneous anisotropic media in which caustics can develop, provided that these media do not allow different reflectors to have identical surface seismic measurements that persist under small perturbations of the reflectors. The latter is essentially the Bolker condition (Guillemin, 1985). This mapping defines map depth-migration, which allows us to go beyond the framework of normal moveout. This normal moveout is at the basis of pre-stack map time-migration making use of RMS velocities, that was treated in the previous sections.

For heterogeneous anisotropic media, the relation between the reflections measured at the surface and the reflectors in the image, is, in the high-frequency regime, governed by ray tracing. We capture this relation schematically with the symbol Λ. Since we assume single-scattering only, this relation is by definition the *canonical relation* of the single scattering modeling or imaging operators that relate the surface seismic measurements to the subsurface image. Map depth-migration essentially determines this canonical relation. Throughout the remainder, we will refer to this relation as *the* canonical relation.

2.5.1 The homogeneous isotropic medium case revisited

To make explicit the connection between the closed-form expressions for pre-stack map time-migration and demigration derived in the previous sections, and the above-mentioned canonical relation, we revisit the isotropic homogeneous medium case. For pure modes (P-P or S-S) in isotropic homogeneous media, the equations determining pre-stack map time-migration using both the horizontal slownesses at the source and receiver, follow directly from the equations for homogeneous VTI media by setting $\frac{dV}{d\theta} = 0$ and $V_{s,r} = V_{P0,S0} = v$. Under these restrictions, equations (2.43), (2.45) and (2.47) thus determine the migrated location $\boldsymbol{x}_m = (x_m, y_m, z_m)$ for homogeneous isotropic media. Because the expressions for the reflector dip in homogeneous VTI media do not contain the derivative of the phase velocity, the equation for the wavenumber associated with the reflector (or the dip covector) $\boldsymbol{\xi}_m$ for homogeneous isotropic media, is given by equation (2.50) with $V_{s,r} = v$. The canonical relation is now formed by collecting the unmigrated and migrated quantities in a table, i.e.,

$$\Lambda' = \{\underbrace{(x_u, y_s, y_r, t_u, -\omega(p_s^x + p_r^x), -\omega p_s^y, -\omega p_r^y, \omega}_{\text{reflection}}; \underbrace{\boldsymbol{x}_m, \boldsymbol{\xi}_m)}_{\text{reflector}}\}, \tag{2.71}$$

where ω is the angular frequency. The prime in Λ' indicates that the canonical relation is here restricted to common azimuth. Hence, equations (2.43), (2.45), (2.47) and (2.50), subjected to the restrictions $\frac{dV}{d\theta} = 0$ and $V_{s,r} = V_{P0,S0} = v$, determine the reflector $(\boldsymbol{x}_m, \boldsymbol{\xi}_m)$ in the image from the reflections in the data $(x_u, y_s, y_r, t_u, -\omega(p_s^x + p_r^x), -\omega p_s^y, -\omega p_r^y, \omega)$, and thus define the canonical relation, restricted to common azimuth (i.e., Λ'), of the single scattering imaging operator in homogeneous isotropic media; these equations evaluate (2.71) 'from left to right'.

To find the equations determining map time-demigration in isotropic homogeneous media, we need to find x_u, $y_{s,r}$, t_u, and $p_{s,r}^{x,y}$ as a function of $(\boldsymbol{x}_m, \boldsymbol{\xi}_m)$. By setting $\frac{dV}{d\theta} = 0$ and $V_{s,r} = v$ in equations (2.61) and (2.63)-(2.65), we get the system of equations that needs to be solved to find the angles $\theta_{s,r}$ and azimuths $\gamma_{s,r}$ at the reflector. For pure modes, the resulting system of equations and the solutions for $(s,r)_\theta = (\sin\theta_s, \sin\theta_r)$ and $(s,r)_\gamma = (\sin\gamma_s, \sin\gamma_r)$, are given in Appendix B. Once $(s,r)_\theta$ and $(s,r)_\gamma$ are found, we calculate the source and receiver locations $(x_u,\ y_{s,r})$, two-way traveltime t_u, and the horizontal slownesses $p_{s,r}^{x,y}$ from

simple geometrical considerations. This gives

$$x_u = x_{s,r} = x_m - \mathrm{sgn}(\nu_x)\frac{z_m\,(s,r)_\theta\,\sqrt{1-(s,r)_\gamma^2}}{\sqrt{1-(s,r)_\theta^2}}\,, \tag{2.72}$$

$$y_{s,r} = y_m + \frac{z_m\,(s,r)_\gamma\,(s,r)_\theta}{\sqrt{1-(s,r)_\theta^2}}\,, \tag{2.73}$$

$$t_u = \frac{z_m}{v}\left(\sqrt{1-s_\theta^2}+\sqrt{1-r_\theta^2}\right)\,, \tag{2.74}$$

$$p^x_{s,r} = -\mathrm{sgn}(\nu_x)\sqrt{1-(s,r)_\gamma^2}\frac{(s,r)_\theta}{v}\,, \tag{2.75}$$

$$p^y_{s,r} = (s,r)_\gamma\frac{(s,r)_\theta}{v}\,. \tag{2.76}$$

These equations determine the reflection in the data $(x_u, y_s, y_r, t_u, -\omega(p_s^x + p_r^x), -\omega p_s^y, -\omega p_r^y, \omega)$ from the reflector in the image $(\boldsymbol{x}_m, \boldsymbol{\xi}_m)$, and thus define the canonical relation, restricted to common azimuth (i.e., Λ'), of the single scattering modeling operator in homogeneous isotropic media; they evaluate (2.71) 'from right to left'. Note that the common-offset equivalent equations to (2.72)-(2.76) are given by equations (2.21)-(2.24).

In the section on map time-migration and demigration in isotropic homogeneous media, we showed that only the slope in common-offset gathers (as opposed to both common-offset and common-midpoint gathers, or common-source and common-receiver gathers) needs to be known, to map the surface seismic measurements, restricted to common offset and common azimuth, to the reflectors in the image. Under these restrictions, the canonical relation for pre-stack map time-migration is therefore given by

$$\Lambda'' = \{\underbrace{(x_u, y_u - h, y_u + h, t_u, -\omega p_u^x, -\omega p_u^y, \omega}_{\text{reflection}}; \underbrace{\boldsymbol{x}_m, \boldsymbol{\xi}_m)}_{\text{reflector}}\}. \tag{2.77}$$

Here, the double prime indicates the restriction of the canonical relation to both common offset and common azimuth. Not counting the angular frequency ω as a variable, there are six variables $(x_u, y_u, t_u, p_u^{x,y}, h)$ describing the data. The number of variables $(\boldsymbol{x}_m, \boldsymbol{\xi}_m)$ describing the reflector is also six. Hence, for common-offset and common-azimuth data, pre-stack map time-migration and demigration are direct mappings from input to output variables.

2.5.2 The canonical relation in heterogeneous anisotropic media

Since we derived all expressions for map time-migration in the rotated coordinate system with the y-axis positive in the source-to-receiver direction, these expressions implicitly assume common azimuth, i.e., $x_r = x_s = x_u$. Lifting this restriction, the canonical relation

is written as

$$\Lambda = \{(\underbrace{\boldsymbol{x}_s^h, \boldsymbol{x}_r^h, t_u, -\omega\boldsymbol{p}_s^h, -\omega\boldsymbol{p}_r^h, \omega}_{\text{reflection}}; \underbrace{\boldsymbol{x}_m, \boldsymbol{\xi}_m}_{\text{reflector}})\}, \tag{2.78}$$

where $\boldsymbol{x}_{s,r}^h = (x_{s,r}, y_{s,r})$ are the source and receiver locations respectively, $t_u = t_s + t_r$ is the two-way traveltime, $\boldsymbol{p}_{s,r}^h = (p_{s,r}^x, p_{s,r}^y)$ are the horizontal slownesses at the sources and receivers, $\boldsymbol{x}_m = (x_m, y_m, z_m)$ is the reflector position in the image, and $\boldsymbol{\xi}_m = \boldsymbol{\xi}_s + \boldsymbol{\xi}_r$ is the wave vector associated with the reflector; here the superscript h denotes horizontal components only. Note that here the vertical components of the slowness vectors $\boldsymbol{p}_{s,r}$ are defined to point in the negative z-direction, as throughout the whole text.

In general media, Λ is evaluated by ray tracing. Writing the solution to the ray-tracing equations, subject to initial conditions $(\boldsymbol{x}_0, \boldsymbol{\xi}_0)$ at time 0, in the general form $(\boldsymbol{x}(\boldsymbol{x}_0, \boldsymbol{\xi}_0, t), \boldsymbol{\xi}(\boldsymbol{x}_0, \boldsymbol{\xi}_0, t))$, the canonical relation becomes, for a horizontal acquisition surface,

$$\begin{aligned}\Lambda \;=\; \Big\{\Big(&\boldsymbol{x}^h(\boldsymbol{x}_m, \boldsymbol{\xi}_s, t_s), \boldsymbol{x}^h(\boldsymbol{x}_m, \boldsymbol{\xi}_r, t_r), t_u, \boldsymbol{\xi}^h(\boldsymbol{x}_m, \boldsymbol{\xi}_s, t_s), \boldsymbol{\xi}^h(\boldsymbol{x}_m, \boldsymbol{\xi}_r, t_r), \omega; \boldsymbol{x}_m, \boldsymbol{\xi}_s + \boldsymbol{\xi}_r\Big) \\ &\text{such that } [z(\boldsymbol{x}_m, \boldsymbol{\xi}_s, t_s) = 0, z(\boldsymbol{x}_m, \boldsymbol{\xi}_r, t_r) = 0]\Big\},\end{aligned}$$

with $\boldsymbol{\xi}^h \equiv (\xi_x, \xi_y)$. This symbolically represents a table parametrized by upward ray-tracing from the reflector at location $\boldsymbol{x}_m$.

2.5.3 Pre-stack map depth-migration and demigration in heterogeneous anisotropic media

Defining $\boldsymbol{u} \equiv (x_s, y_s, x_r, y_r, t_u)$ and $\boldsymbol{v} \equiv (\omega p_s^x, \omega p_s^y, \omega p_r^x, \omega p_r^y, \omega)$, $(\boldsymbol{u}, \boldsymbol{v})$ is a surface seismic measurement characterized by the source and receiver locations, two-way traveltime, and slopes in common source and receiver gathers; it is an element of a phase space U. Similarly, defining $\boldsymbol{m} \equiv \boldsymbol{x}_m$ and $\boldsymbol{\mu} \equiv \boldsymbol{\xi}_m$, $(\boldsymbol{m}, \boldsymbol{\mu})$ describes a reflector defined by its location and orientation in the image; it is an element of a phase space M.

Let P_U now denote the projection of Λ on U and P_M denote the projection of Λ on M. These projections extract, respectively, the surface seismic reflections $(\boldsymbol{u}, \boldsymbol{v})$ and the reflectors in the image $(\boldsymbol{m}, \boldsymbol{\mu})$ from the canonical relation Λ, i.e., the table evaluated using ray tracing. Guillemin (1985), in his paper analyzing the generalized Radon transform, introduced the Bolker condition on the canonical relation, which basically states that the medium does not allow different reflectors to have identical surface seismic measurements that persist in being identical under small perturbations of the reflectors, *given a velocity model and given an acquisition geometry.* The Bolker condition is crucial in the development of seismic inverse scattering theory in the presence of caustics (Stolk & De Hoop, 2002a). Here, we explain that this condition can be understood in terms of map migration.

In our application, the Bolker condition reduces to the condition that $P_U : \Lambda \to P_U(\Lambda)$ is one-to-one. But then, we can introduce the mapping $P_M \circ P_U^{-1} : (\boldsymbol{u}, \boldsymbol{v}) \mapsto (\boldsymbol{m}, \boldsymbol{\mu})$, which is precisely map depth-migration; here $\circ$ denotes composition, which can be thought of as a cascade of mappings (i.e., P_U^{-1} followed by P_M). Hence, the Bolker condition is the *condition of applicability* of map depth-migration in heterogeneous anisotropic media

allowing the formation of caustics, given a velocity model and acquisition geometry. In other words, if, given a velocity model and acquisition geometry, you can map depth-migrate *without ambiguity* in either the migrated location or orientation, then the Bolker condition is simply satisfied.

Map depth-demigration does not require the Bolker condition. To find the surface seismic measurements from a reflector in the image through map depth-demigration, we need to specify the scattering angle (for both 2D and 3D seismics) and azimuth (only for 3D) at the reflector, in addition to the reflector location and orientation, i.e., $(\boldsymbol{m}, \boldsymbol{\mu})$. We introduce these (angle) coordinates by parameterizing the subsets $\{(\boldsymbol{m}, \boldsymbol{\mu})\} = \text{const}$ on Λ, and denote them by $\boldsymbol{e}$. Thus $(\boldsymbol{m}, \boldsymbol{\mu}, \boldsymbol{e})$ form local coordinates on Λ. Map depth-demigration then follows from the mapping $(\boldsymbol{m}, \boldsymbol{\mu}, \boldsymbol{e}) \mapsto (\boldsymbol{u}, \boldsymbol{v})$. This mapping was denoted Ω in Stolk & De Hoop (2002a). In the absence of caustics, $\boldsymbol{e}$ can be chosen to be acquisition offset and azimuth (as we did in the section on pre-stack map time-demigration in homogeneous VTI media).

2.6 Discussion

We have presented closed-form 3D pre-stack map time-migration expressions for qP-qP, qP-qSV, and qSV-qSV waves in homogeneous VTI media, that specialize to the expressions for P-P, P-S, and S-S waves in homogeneous isotropic media. In addition, we have presented closed-form expressions for pre-stack map time-migration and demigration in the common-offset domain for pure mode (P-P or S-S) waves in homogeneous isotropic media, that use only the slope in the common-offset domain. This provides an additional advantage over methods where both p_u and p_h (or equivalently both p_s and p_r, the slopes at the source and receiver position) are required, especially since estimating slopes can be cumbersome in the presence of noise. All the derived pre-stack expressions reduce properly to their zero-offset counterparts. Our closed-form expressions for pre-stack map time-migration can be exploited in existing velocity-inversion algorithms that use map migration in such media. In particular, our expressions for pre-stack map time-migration of qP waves in homogeneous VTI media, can be used to determine the anellipticity parameter η (and the zero-dip NMO velocity $V_{NMO}(0)$) for such media in a time-migration velocity analysis context. Note that for media with mild lateral and vertical velocity variations, our equations can be used provided the velocity is replaced by the local RMS velocity. The kinematic equivalence of TI media and orthorhombic media in the symmetry planes, generalizes our results for VTI media to these planes in orthorhombic media. In addition, our expressions for VTI media are applicable to the vertical symmetry plane of TTI homogeneous media, which contains the symmetry axis.

Not surprisingly, our closed-form expressions for pre-stack map time-migration in homogeneous isotropic and VTI media, exemplify that, for such media, given an offset and azimuth, the mapping from the surface seismic measurements (i.e., the source and receiver locations, two-way traveltime, *and* slopes in common source and receiver gathers) to the reflectors in the image (i.e., location *and* orientation), is one-to-one. We explained that the condition of applicability of pre-stack map depth-migration in heterogeneous anisotropic

media that allow the formation of caustics - i.e., that, given a velocity model and acquisition geometry, you can map depth-migrate *without ambiguity* in both the migrated location and orientation - coincides with the Bolker condition for seismic inverse scattering. In addition, we have shown that for homogeneous media, our pre-stack map time-migration and demigration expressions define the canonical relation of the single scattering modeling and imaging operators in such media.

The tangential directions to the recorded wavefronts in seismic data, are the directions, locally, in which the data are smooth. Imaging these wavefronts provides the directions in which the medium perturbations are smooth. In mathematical terms, these smooth directions are the wavefront sets of the seismic data and the medium perturbations. In the context of sparsely representing the data and the image, most sparseness (or compression) will be accomplished in the smooth directions, i.e., along the wavefront sets of the data and the medium perturbations. Since, given a scattering angle and azimuth, map migration provides a one-to-one mapping from the singular directions in the data (i.e., the directions normal to the wavefronts) to the singular directions in the image (i.e., the normals to the reflectors), one can thus think of map migration (and demigration) as a mapping between optimally sparse directions. Therefore, it seems that map migration (or demigration) is a suitable vehicle to sparsely represent the imaging (or modeling) operator.

Chapter 3

Explicit expressions for map time-migration with qP-waves in weakly anisotropic VTI media[1]

3.1 Summary

In the context of time migration, map migration achieves a one-to-one correspondence between the reflections recorded at the surface (position, time, and slopes) and the reflectors in the subsurface (position and orientation), through explicit use of the slopes in the data. Previously, Douma & de Hoop (2006) (see Chapter 2) presented pre-stack map time-migration equations that are valid in homogeneous transversely isotropic with a vertical symmetry axis and for arbitrary levels of anisotropy. As pointed out by Thomsen (1986), however, many rocks exhibit only weak anisotropy. Here, I apply the weak-anisotropy approximation to the post-stack and pre-stack map time-migration equations for qP-waves in transversely isotropic media with a vertical symmetry axis. These expressions depend on only the anellipticity parameter η and the zero-dip NMO velocity $V_{NMO}(0)$, and can be used to estimate η in the context of time-migration velocity analysis. The accuracy of the equations is verified with numerical examples, which also quantify the error in time imaging when the anellipticity is ignored. That error indicates that accurately accounting for anisotropy is highly relevant to positioning of boreholes.

3.2 Introduction

Imaging the subsurface involves mapping seismic waves measured at the earth's surface down to the subsurface locations from which the waves reflected (once). Under certain conditions (Guillemin, 1985; ten Kroode *et al.*, 1998; de Hoop & Brandsberg-Dahl, 2000; Stolk & De Hoop, 2002b; Douma & de Hoop, 2006), for arbitrary complex and anisotropic media, the slope information and the position (in space and time) of the reflections are related through a one-to-one mapping to the subsurface position and the orientation of the reflectors. This mapping is referred to as (pre-stack) map migration (Douma & de Hoop, 2006, or Chapter 2 of this thesis). In the context of time imaging, these conditions are simply always satisfied, and the mapping is thus always one-to-one. Although most imaging algorithms use the slope information in the data only implicitly, map migration uses this

[1]This chapter, along with Appendix E, will be submitted for publication as: Huub Douma, 2006, Explicit expressions for map time-migration with qP-waves in weakly anisotropic VTI media.

information explicitly. Because of this one-to-one mapping property of map migration, it has had many applications in velocity estimation (Gjoystdal & Ursin, 1981; Gray & Golden, 1983; Maher *et al.*, 1987; Sword, 1987; Iversen & Gjoystdal, 1996; Billette & Lambaré, 1998; Iversen *et al.*, 2000).

Douma & de Hoop (2006) (see Chapter 2) derived closed-form expressions for map time-migration in transversely isotropic (TI) media with a vertical symmetry axis (VTI). Through careful numerical analysis, Alkhalifah & Tsvankin (1995) have shown that the time signatures (e.g., reflection moveout, dip moveout, and post-stack and pre-stack time-migration operators) of qP-waves in such media are governed mainly by the anellipticity parameter η and the zero-dip NMO velocity, $V_{NMO}(0)$. In addition, Tsvankin (1996) shows numerically that the influence of the vertical shear-wave velocity V_{S0} on the kinematics of qP-waves in TI media is negligible, while Alkhalifah (1998) shows analytically that all time-related processing in such media depends exactly on η and $V_{NMO}(0)$ only, when $V_{S0} = 0$ km/s. The expressions for pre-stack map time-migration of qP-waves presented in Chapter 2, however, depend on the four parameters η, $V_{NMO}(0)$, V_{P0}, and V_{S0}. In principle it should be possible to show the explicit dependence of these equations on η and $V_{NMO}(0)$ only, through use of the acoustic approximation of Alkhalifah (1998) (i.e., setting $V_{S0} = 0$ km/s). In Chapter 2, however, this approximation is invoked only to simplify the presented nonlinear system of equations for pre-stack map time-demigration in VTI media (see Appendix A), and to show explicitly the sole dependence of this system on η and $V_{NMO}(0)$.

Many rocks in the earth exhibit only weak anisotropy (Thomsen, 1986). Therefore, in this paper, I apply the weak-anisotropy approximation to derive closed-form expressions for map time-migration of qP-waves in weakly anisotropic VTI media. As it turns out, by invoking the weak-anisotropy approximation the resulting expressions for map time-migration of qP-waves in VTI media are dependent on η and $V_{NMO}(0)$ only. The equations can be used to invert for the anellipticity parameter η in the context of time-migration velocity analysis.

The outline of this paper is as follows. First, to establish a connection with earlier work on time-signatures of qP-waves in VTI media, I show that the expressions for the zero-offset migrated traveltime and position presented in Chapter 2 are equivalent to the expressions derived by Cohen (1998). Then, I proceed to derive the weak-anisotropy approximations for post-stack (or, equivalently, zero-offset) time migration and demigration respectively, followed by the weak-anisotropy approximations for pre-stack (or finite-offset) map migration. The accuracy of all presented equations is verified using numerical examples, which also quantify the migration error when anisotropy is ignored.

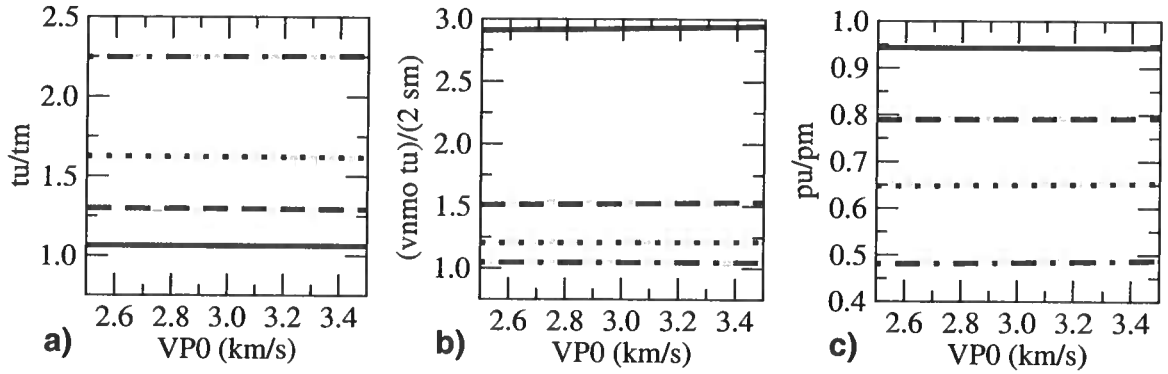


Figure 3.1. Dependence on V_{P0} of the ratios (a) t_u/t_m, (b) $V_{NMO}(0)t_u/(2s_m)$ with $s_m := x_u - x_m$, and (c) p_u/p_m. All ratios are calculated for the values $p_u = 0.10$ s/km (drawn lines), $p_u = 0.18$ s/km (dashed lines), $p_u = 0.22$ s/km (dotted lines) and $p_u = 0.25$ s/km (dashed-dotted lines) (i.e., increasing slopes in the data), with the values $V_{NMO}(0) = 3.29$ km/s, $\eta = 0.0833$, and $V_{P0}/V_{S0} = 3/2$ (i.e., $f = 5/9$).

3.3 Post-stack map migration

From Douma & de Hoop (2006), the zero-offset expressions for 3D map migration in homogeneous VTI media [see Chapter 2, equations (2.55)-(2.58)] are given by

$$t_m = \frac{V(\theta)t_u}{V_{P0}}\left(\sqrt{1-V^2(\theta)p_u^2} - p_u \left.\frac{dV}{d\theta}\right|_\theta\right), \tag{3.1}$$

$$(x,y)_m = (x,y)_u - \frac{V^2(\theta)p_u^{x,y}t_u}{2} - \frac{V(\theta)p_u^{x,y}t_u}{2}\left.\frac{dV}{d\theta}\right|_\theta \sqrt{\frac{1}{V^2(\theta)p_u^2} - 1}, \tag{3.2}$$

$$p_m^{x,y} = \frac{V(\theta)p_u^{x,y}}{V_{P0}\sqrt{1-V^2(\theta)p_u^2}}, \tag{3.3}$$

where $t_{u,m}$ are the unmigrated and migrated two-way traveltimes, $(x,y)_{u,m}$ are the unmigrated and migrated horizontal locations, $p_u^{x,y} := (1/2)\partial t_u/\partial(x,y)_u$ are the slopes in the x- and y-directions in the data, $p_u := \sqrt{(p_u^x)^2 + (p_u^y)^2}$, $p_m^{x,y} := (1/2)\partial t_m/\partial(x,y)_m$ are the slopes in the x- and y-directions in the image, V_{P0} is the vertical P-wave velocity, and $V(\theta)$ is the phase velocity at angle θ with the symmetry axis. The notation (x,y) indicates x or y. Note that the subscripts u and m simply denote unmigrated and migrated variables, respectively. Equations (3.1) - (3.3) thus determine the reflector position and orientation from the unmigrated quantities associated with a reflection in the data. The expressions for 2D map migration can be found from the 3D expressions by setting $x_m = x_u = 0$ and $p_m^x = p_u^x = 0$. To establish the connection with earlier work on map migration in homo-

geneous VTI media, Appendix E shows that equations (3.1) and (3.2), respectively, are equivalent to the expressions derived by Cohen (1998, his equation (28)) for the migrated vertical two-way traveltime and the horizontal migrated location.

Recall that Alkhalifah & Tsvankin (1995) showed that qP-wave time signatures in (laterally homogeneous) VTI media are governed mainly by the anellipticity parameter η and the zero-dip NMO-velocity $V_{NMO}(0)$, and are largely independent of V_{P0}. Here, I confirm this independence in the context of post-stack map time-migration in Figure 3.1. The values $\eta = 0.0833$, $V_{NMO}(0) = 3.29$ km/s, and the horizontal slowness components in this figure are the same as those used by Alkhalifah & Tsvankin (1995, Figure 17). However, since in their paper the value used for the vertical shear-wave velocity V_{S0} is omitted, Figure 3.1a closely resembles Alkhalifah and Tsvankin's Figure 17, but is not identical. Figure 3.1 shows results for $V_{S0} = 0$ (grey) and $V_{P0}/V_{S0} = 1.5$ (black), which in practice can be treated as a lower bound on the ratio of P- and S-wave velocity. As indicated by the black curves, the zero-offset map time-migration equations are indeed governed mainly by η and $V_{NMO}(0)$ and are largely independent of V_{P0}. In addition, the grey curves confirm that qP-wave zero-offset map time-migration is completely independent of V_{P0} if V_{S0} is set to zero (Alkhalifah, 1998).

Under the weak-anisotropy approximation, i.e., assuming that the Thomsen parameters $(\delta, \epsilon) \ll 1$, the expressions for the phase-velocity and its derivative reduce to [e.g. (Tsvankin, 2001, p. 23)]

$$V(\theta) \approx V_{P0}\left(1 + \delta \sin^2\theta + (\epsilon - \delta)\sin^4\theta\right) , \tag{3.4}$$

and

$$\frac{dV}{d\theta} \approx 2V_{P0}\sin\theta\cos\theta\left[\epsilon - (\epsilon - \delta)\left(1 - 2\sin^2\theta\right)\right] . \tag{3.5}$$

Using equations (3.4) and (3.5) in (3.1) and linearizing in δ and ϵ, I find

$$t_m \approx t_u\sqrt{1 - p_u^2V_{P0}^2\left(1 + 2\delta\right) - 6\left(\epsilon - \delta\right)p_u^4V_{P0}^4 + 4\left(\epsilon - \delta\right)p_u^6V_{P0}^6} , \tag{3.6}$$

where I used the approximation $\sin\theta \approx p_uV_{P0}$ whenever $\sin\theta$ was multiplied with δ or ϵ, in order to consistently linearize in the Thomsen parameters. Since in the weak-anisotropy approximation we have $\eta = (\epsilon - \delta)/(1 + 2\delta) \approx \epsilon - \delta$ and since in general we have $V_{NMO}(0) = V_{P0}\sqrt{1 + 2\delta}$, we can rewrite equation (3.6) to find

$$t_m \approx t_u\sqrt{1 - p_u^2V_{NMO}^2(0) - 4\eta p_u^4V_{NMO}^4(0)\left(\frac{3}{2} - p_u^2V_{NMO}^2(0)\right)} . \tag{3.7}$$

Similarly, by using equations (3.4) and (3.5) in equations (3.2) and (3.3) and further

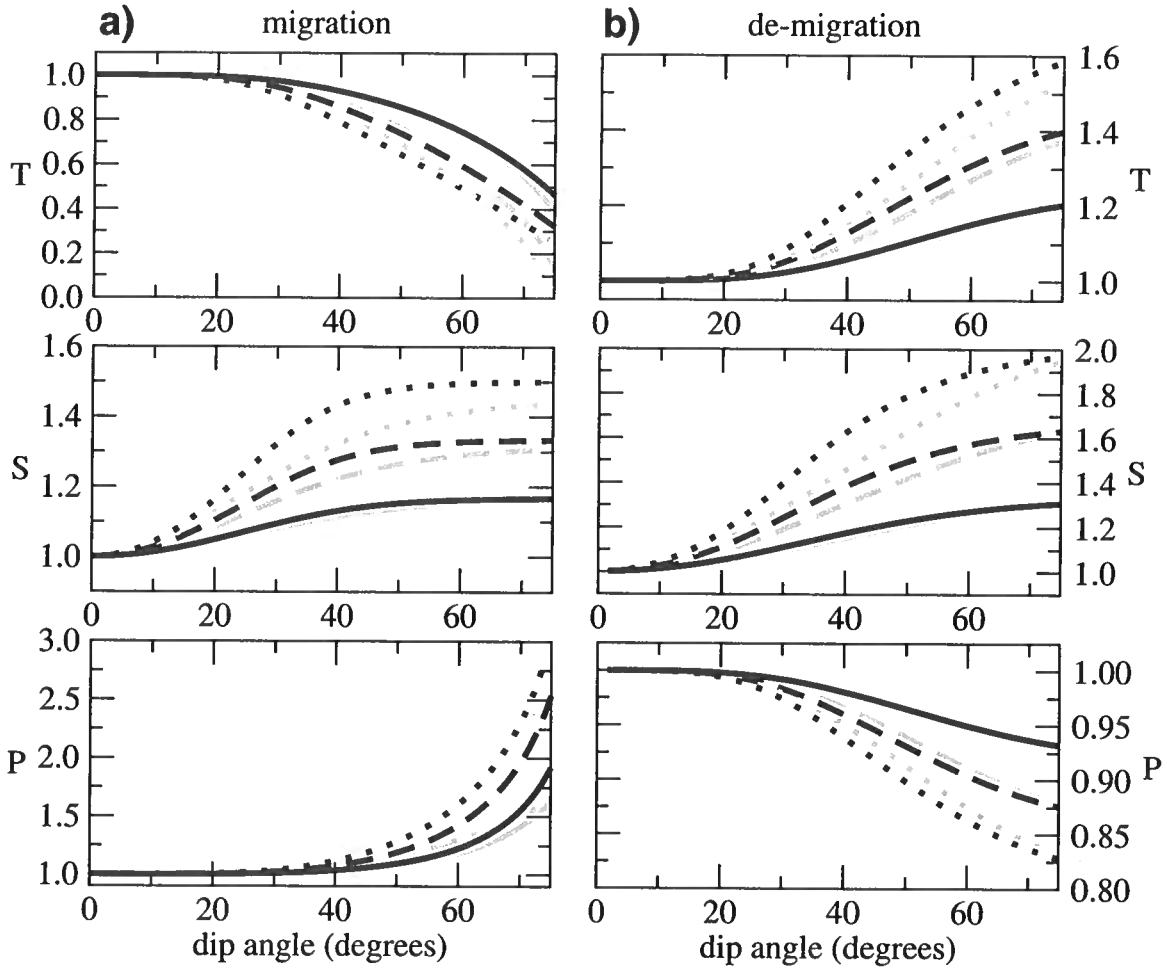


Figure 3.2. a) Ratios (from the top down) $t_m/(t_m|_{\eta=0})$, $s_m/(s_m|_{\eta=0})$ and $p_m/\ (p_m|_{\eta=0})$ as a function of reflector dip, using the weak-anisotropy approximated equations (grey) and the exact expressions (black). b) Ratios (from the top down) $t_u/(t_u|_{\eta=0})$, $s_u/(s_u|_{\eta=0})$ with $s_u := x_u - x_m$, and $p_u/(p_u|_{\eta=0})$, as a function of reflector dip using the weak-anisotropy approximated equations (grey) and the exact expressions (black). All values are calculated with $V_{NMO}(0)$ and f as in Figure 3.1, $V_{P0} = 3$ km/s, $\eta = 0.0833$ (solid), $\eta = 0.1666$ (dashed), and $\eta = 0.25$ (dotted).

linearizing in the Thomsen parameters, I find

$$(x,y)_m \approx (x,y)_u - \frac{V_{NMO}^2(0)p_u^{x,y}t_u}{2}\left(1+4\eta p_u^2 V_{NMO}^2(0)\left[1-\frac{p_u^2 V_{NMO}^2(0)}{2}\right]\right), \quad (3.8)$$

$$p_m^{x,y} \approx \frac{p_u^{x,y}}{\sqrt{1-p_u^2V_{NMO}^2(0)\left(1+2\eta p_u^2 V_{NMO}^2(0)\right)}}. \quad (3.9)$$

Indeed, in the weak-anisotropy approximation, the dependence of the map time-migration equations for qP-waves in homogeneous VTI media on V_{P0} disappears; also for $\eta = 0$ and $V_{NMO}(0) = V_{P0} = v$ the equations indeed reduce to their isotropic equivalents [see Chapter 2 for the isotropic equations, i.e., equations (2.17)-(2.19)].

The accuracy of equations (3.7)-(3.9) is verified in Figure 3.2a. From the top downwards, this figure shows the migrated time, horizontal displacement (i.e. $s_m := x_u - x_m$), and migrated horizontal slowness, normalized by their elliptic counterparts [i.e., $t_m/(t_m|_{\eta=0})$, $s_m/(s_m|_{\eta=0})$ and $p_m/(p_m|_{\eta=0})$, respectively]. The grey lines show the values resulting from the weak-anisotropy approximations, which are indeed close to their true values (black lines) for $V_{NMO}(0) = 3.29$ km/s, $V_{P0} = 3$ km/s, $f = \frac{5}{9}$ (i.e., $V_{P0}/V_{S0} = \frac{3}{2}$) and $\eta = 0.0833$ (drawn lines), but deviate from the exact values for larger values of η [i.e., $\eta = 0.1666$ (dashed) and $\eta = 0.25$ (dotted)]. Note that overall the weak-anisotropy approximation becomes worse for larger angles, but is highly accurate for the position (horizontal and in time) and orientation up to about 60 degree dips for $\eta = 0.0833$. This decrease in accuracy of the weak-anisotropy approximations with increasing dip is due to the fact that higher orders of $\sin\theta$ usually go together with higher orders in the Thomsen parameters, which are ignored in the weak-anisotropy approximation.

Because Figure 3.2a shows the map time-migrated quantities normalized by their elliptical counterparts, this figure also indicates the errors in ignoring the anellipticity parameter η in post-stack time-imaging; i.e., the amount of deviation from unity for each quantity indicates the error due to ignoring anisotropy. I normalize by the elliptic counterparts (i.e., $\eta = 0$), because for elliptical anisotropy all time-domain processing is kinematically equivalent to the isotropic case (e.g., Tsvankin (2001)); i.e., ignoring η for qP-waves in (laterally homogeneous) VTI media is the same as ignoring the presence of anisotropy all together. Clearly, ignoring the presence of anisotropy leads to significant errors in the migrated location (i.e., s_m and t_m) and orientation of the reflector (i.e., p_m) since the ratios of the migrated quantities x_m, t_m, and p_m to their elliptic equivalents are substantially different from unity, even for moderate dips. The mispositioning in time and orientation of the reflector, as a result of ignoring η, however, is negligible for dips up to about 20 degrees for $\eta = 0.0833$. In contrast, the mispositioning in horizontal position is in this case negligible only up to about 10 degrees. Even though we are dealing with homogeneous media here, accurately accounting for the presence of anisotropy seems highly relevent for determining the position of a borehole in heterogeneous media. Mis-positioning in post-stack imaging as a result of ignoring subsurface anisotropy has been studied by Larner & Cohen (1993) and Alkhalifah & Larner (1994).

3.4 Post-stack map demigration

The post-stack map time-demigration equations in homogeneous VTI media are given by [see Chapter 2, equations (2.68)-(2.67)]

$$t_u = \frac{t_m\sqrt{1+V_{P0}^2p_m^2}}{\left(\frac{V(\theta)}{V_{P0}} - p_m\left.\frac{dV}{d\theta}\right|_\theta\right)}, \tag{3.10}$$

$$(x,y)_u = (x,y)_m + \frac{V_{P0}^2p_m^{x,y}t_m\left(V(\theta)+\frac{1}{V_{P0}p_m}\left.\frac{dV}{d\theta}\right|_\theta\right)}{2\left(V(\theta)-V_{P0}p_m\left.\frac{dV}{d\theta}\right|_\theta\right)}, \tag{3.11}$$

$$p_u^{x,y} = \frac{V_{P0}p_m^{x,y}}{V(\theta)\sqrt{1+V_{P0}^2p_m^2}}, \tag{3.12}$$

with $p_m := \sqrt{(p_m^x)^2+(p_m^y)^2}$. Then, using equations (3.4) and (3.5) and again consistently linearizing in δ and ϵ, I find

$$t_u \approx t_m\sqrt{1+p_m^2V_{NMO}^2(0)+6\eta\frac{p_m^4V_{NMO}^4(0)}{1+p_m^2V_{NMO}^2(0)}}, \tag{3.13}$$

$$(x,y)_u \approx (x,y)_m + \frac{V_{NMO}^2(0)p_m^{x,y}t_m}{2}\left(1+4\eta\frac{p_m^2V_{NMO}^2(0)}{1+p_m^2V_{NMO}^2(0)}\right), \tag{3.14}$$

$$p_u^{x,y} \approx p_m^{x,y}\sqrt{1+p_m^2V_{NMO}^2(0)+2\eta\frac{p_m^4V_{NMO}^4(0)}{1+p_m^2V_{NMO}^2(0)}}. \tag{3.15}$$

For $\eta = 0$ and $V_{NMO}(0) = V_{P0} = v$, these approximations reduce to the isotropic expressions [see Chapter 2, equations (2.28)-(2.30)]. To obtain these expressions for the demigration problem, I use the exact expression $\sin\theta = p_mV_{P0}/\sqrt{1+p_m^2V_{P0}^2}$ (which follows from writing $p_u = \sin\theta/V(\theta)$ in the 2D equivalent expression of equation (3.12) and subsequently solving for $\sin\theta$).

To assess the accuracy of the approximations, Figure 3.2b shows the normalized (again by their elliptic equivalents) 2D de-migrated quantities calculated with the exact and approximated equations, for the same values of η, $V_{NMO}(0)$, V_{P0}, and f as used in Figure 3.2a; i.e., this figure shows (from the top down-wards) $t_u/\ t_u|_{\eta=0}$, $s_u/\ s_u|_{\eta=0}$ with $s_u := x_u - x_m$, and $p_u/\ p_u|_{\eta=0}$. Again, the approximated values (grey) are in close agreement with the true values (black) for $\eta = 0.833$, but the accuracy of the approximated quantities decreases with increasing dip and η. Note that in the derivation of the weak-anisotropy approximations to the demigration equations, it is not necessary to make the additional approximation $\sin\theta \approx p_uV_{P0}$ whenever $\sin\theta$ is multiplied with a Thomsen parameter, as was necessary in the derivation of the approximations to the migration equations. Instead, we can just

use the exact expression $\sin\theta = p_m V_{P0}/\sqrt{1+p_m^2 V_{P0}^2}$. This explains the relatively higher accuracy of the unmigrated horizontal slowness even at large dips and large values of η, when compared to the accuracy obtained in the weak-anisotropy approximation to p_m (cf. Figure 3.2a, bottom subfigure).

At first sight, the reader might be confused by the difference between the middle subfigures of Figures 3.2a and b, since both subfigures show the quantity $(x_u - x_m)/\left.(x_u - x_m)\right|_{\eta=0}$. However, in the migration case (Figure 3.2a) x_m is varying while x_u is fixed, wheras in the demigration case (Figure 3.2b) x_u is varying while x_m is fixed. This explains the difference between both subfigures.

3.5 Pre-stack map migration

For pre-stack (or finite-offset) map migration in homogeneous VTI media, the equations for the reflector location and orientation in the subsurface are given by [see Chapter 2, equations (2.43), (2.45), and (2.52)]

$$t_m = \frac{2t_u}{V_{P0}}\left(\frac{1}{V_s\left(\sqrt{1-V_s^2p_s^2}-p_s\left.\frac{dV}{d\theta}\right|_s\right)}+\frac{1}{V_r\left(\sqrt{1-V_r^2p_r^2}-p_r\left.\frac{dV}{d\theta}\right|_r\right)}\right)^{-1} \tag{3.16}$$

$$(x,y)_m = (x,y)_{s,r} - t_u p_{s,r}^{x,y}\left(V_{s,r}+\sqrt{\frac{1}{V_{s,r}^2p_{s,r}^2}-1}\left.\frac{dV}{d\theta}\right|_{s,r}\right) \times \frac{\left(\sqrt{1-V_{r,s}^2p_{r,s}^2}-p_{r,s}\left.\frac{dV}{d\theta}\right|_{r,s}\right)}{\frac{1}{V_r}\left(\sqrt{1-V_s^2p_s^2}-p_s\left.\frac{dV}{d\theta}\right|_s\right)+\frac{1}{V_s}\left(\sqrt{1-V_r^2p_r^2}-p_r\left.\frac{dV}{d\theta}\right|_r\right)}, \tag{3.17}$$

$$p_m^{x,y} = \frac{1}{V_{P0}}\frac{p_s^{x,y}+p_r^{x,y}}{\sqrt{\frac{1}{V_s^2}-p_s^2}+\sqrt{\frac{1}{V_r^2}-p_r^2}}, \tag{3.18}$$

where $V_{s,r}$ are the phase-velocities at the source and receiver, $p_{s,r}$ are the horizontal slownesses at the source and receiver (i.e., $p_{s,r} := \sqrt{(p_{s,r}^x)^2+(p_{s,r}^y)^2}$ with $p_{s,r}^{x,y}$ the components of $p_{s,r}$ in the x and y directions), and h is the half-offset. Note the order of the subscripts s,r and r,s in equation (3.17). That is, choosing $p_{s,r}$ to be p_s goes together with letting $p_{r,s}$ be p_r. The same holds for $V_{s,r}$ and $V_{r,s}$ and the derivatives.

Comparison of equation (3.16) with equation (3.1) reveals that the right hand side of equation (3.16) involves the sum of two fractions with denominators that have the same form as $V_{P0}t_m/t_u$ in equation (3.1). Using the form of the weak-anisotropy approximation to t_m for post-stack map time-migration [i.e., equation (3.7)] in order to approximate both

fractions in equation (3.16), it follows that the weak-anisotropy approximation to t_m for pre-stack map migration is given by

$$t_m \approx 2t_u \left(\sum_{i=s,r} \frac{1}{\sqrt{1 - p_i^2 V_{NMO}^2(0) - 4\eta p_i^4 V_{NMO}^4(0) \left\{ \frac{3}{2} - p_i^2 V_{NMO}^2(0) \right\}}} \right)^{-1} . \tag{3.19}$$

Inspection of equations (3.16) and (3.17) reveals that equation (3.17) can be rewritten as

$$(x,y)_m = (x,y)_{s,r} - \frac{V_{P0} p_{s,r}^{x,y} t_m}{2} \frac{V_{s,r}^2 + \sqrt{\frac{1}{p_{s,r}^2} - V_{s,r}^2}\, \frac{dV}{d\theta}\Big|_{s,r}}{V_{s,r} \left(\sqrt{1 - V_{s,r}^2 p_{s,r}^2} - p_{s,r} \frac{dV}{d\theta}\Big|_{s,r} \right)} . \tag{3.20}$$

Since we already have a weak-anisotropy approximation for t_m (cf. equation (3.19)), we need only consider the approximation of the fraction

$$A := \frac{V_{s,r}^2 + \sqrt{\frac{1}{p_{s,r}^2} - V_{s,r}^2}\, \frac{dV}{d\theta}\Big|_{s,r}}{V_{s,r} \left(\sqrt{1 - V_{s,r}^2 p_{s,r}^2} - p_{s,r} \frac{dV}{d\theta}\Big|_{s,r} \right)} , \tag{3.21}$$

to find the approximated expression for $(x,y)_m$. The denominator in this fraction is again of the same form as $V_{P0} t_m / t_u$ in equation (3.1) with its weak-anisotropy approximation given in equation (3.7). We thus have

$$V_{s,r} \left(\sqrt{1 - V_{s,r}^2 p_{s,r}^2} - p_{s,r} \frac{dV}{d\theta}\Big|_{s,r} \right) \approx$$
$$V_{P0} \sqrt{1 - p_{s,r}^2 V_{NMO}^2(0) - 4\eta p_{s,r}^4 V_{NMO}^4(0) \left\{ \frac{3}{2} - p_{s,r}^2 V_{NMO}^2(0) \right\}} . \tag{3.22}$$

Hence, we need only look for a weak-anisotropy approximation to the numerator of A. Using equations (3.4) and (3.5) and consistently linearizing in the Thomsen parameters gives

$$V_{s,r}^2 + \sqrt{\frac{1}{p_{s,r}^2} - V_{s,r}^2}\, \frac{dV}{d\theta}\Big|_{s,r} \approx V_{NMO}^2(0) \left\{ 1 + 4\eta p_{s,r}^2 V_{NMO}^2(0) \left[1 - \frac{p_{s,r}^2 V_{NMO}^2(0)}{2} \right] \right\} . \tag{3.23}$$

Then, using this approximation in equation (3.21) together with equations (3.22) we get

$$A \approx \frac{V_{NMO}^2(0) \left\{ 1 + 4\eta p_{s,r}^2 V_{NMO}^2(0) \left[1 - \frac{p_{s,r}^2 V_{NMO}^2(0)}{2} \right] \right\}}{V_{P0} \sqrt{1 - p_{s,r}^2 V_{NMO}^2(0) - 4\eta p_{s,r}^4 V_{NMO}^4(0) \left\{ \frac{3}{2} - p_{s,r}^2 V_{NMO}^2(0) \right\}}} . \tag{3.24}$$

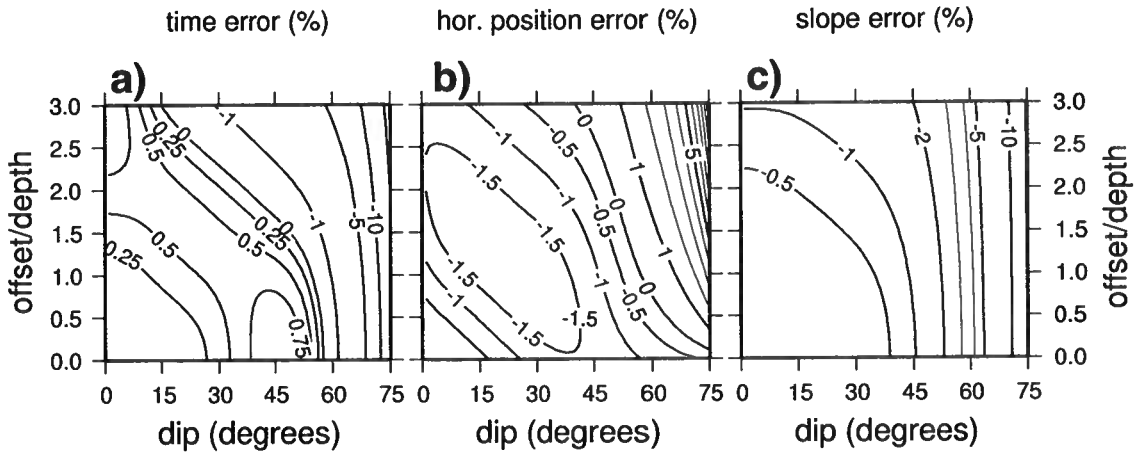


Figure 3.3. Relative errors of weak-anisotropy approximated values with respect to the exact values for (a) migrated time $[(t_m^{weak} - t_m^{exact})/t_m^{exact}]$, (b) horizontal position $[(s_m^{weak} - s_m^{exact})/s_m^{exact}]$ and (c) migrated horizontal slowness $[(p_m^{weak} - p_m^{exact})/p_m^{exact}]$ for 2D pre-stack map migration, as a function of the reflector dip and the offset-to-depth ratio. The values of the relevant parameters used to calculate this figure are $\eta = 0.0833$, $t_m = 2$ s, with $V_{NMO}(0)$, V_{P0} and f the same as in Figure 3.1.

Finally, using this expression together with equation (3.19) in equation (3.20) gives the weak-anisotropy approximation to $(x, y)_m$, i.e.,

$$(x,y)_m \approx (x,y)_{s,r} - p_{s,r}^{x,y} t_u V_{NMO}^2(0) \left\{ 1 + 4\eta p_{s,r}^2 V_{NMO}^2(0) \left[1 - \frac{p_{s,r}^2 V_{NMO}^2(0)}{2} \right] \right\} \times \frac{\sqrt{1 - p_{r,s}^2 V_{NMO}^2(0) - 4\eta p_{r,s}^4 V_{NMO}^4(0) \left\{ \frac{3}{2} - p_{r,s}^2 V_{NMO}^2(0) \right\}}}{\sum_{i=s,r} \sqrt{1 - p_i^2 V_{NMO}^2(0) - 4\eta p_i^4 V_{NMO}^4(0) \left\{ \frac{3}{2} - p_i^2 V_{NMO}^2(0) \right\}}} . \tag{3.25}$$

Note the order of the subscripts s, r and r, s, just as in equation (3.17),

Using the weak-anisotropy approximation in equation (3.18) and again consistently linearizing in δ and ϵ, gives the weak-anisotropy approximation for the horizontal slowness components in the image,

$$p_m^{x,y} \approx \frac{p_s^{x,y} + p_r^{x,y}}{\sum_{i=s,r} \sqrt{1 - p_i^2 V_{NMO}^2(0) \left(1 + 2\eta p_i^2 V_{NMO}^2(0)\right)}} . \tag{3.26}$$

Note that indeed all approximations to the migrated quantities for pre-stack migration reduce to their post-stack equivalent expressions if we set $p_s^{x,y} = p_r^{x,y} = p_u^{x,y}$ [cf. equations

(3.7), (3.8), and (3.9)].

Figure 3.3 shows the relative accuracy of the approximations for pre-stack map time-migration when compared to the exact expressions, as a function of the reflector dip and the offet-to-depth ratio; i.e., Figure 3.3a shows a contour plot of $(t_m^{weak} - t_m^{exact})/t_m^{exact}$, while Figures 3.3b and c show, respectively, contour plots of $(s_m^{weak} - s_m^{exact})/s_m^{exact}$ and $(p_m^{weak} - p_m^{exact})/p_m^{exact}$, all as a function of the reflector dip and the offset-to-depth ratio. Figure 3.3 was calculated by first solving the nonlinear system for 2D pre-stack map demigration in VTI media. This system consists of two nonlinear equations with two unknowns (i.e., the phase angles at the source and receiver); see Chapter 2, equations (2.63) and (2.65), i.e.,

$$\frac{4h}{V_{P0}t_m} = \left(\frac{\left(r_\theta + \frac{\sqrt{1-r_\theta^2}}{V_r} \left.\frac{dV}{d\theta}\right|_r \right)}{\sqrt{1-r_\theta^2} - \frac{r_\theta}{V_r} \left.\frac{dV}{d\theta}\right|_r} - \frac{\left(s_\theta + \frac{\sqrt{1-s_\theta^2}}{V_s} \left.\frac{dV}{d\theta}\right|_s \right)}{\sqrt{1-s_\theta^2} - \frac{s_\theta}{V_s} \left.\frac{dV}{d\theta}\right|_s} \right), \tag{3.27}$$

and

$$\nu = -\left(\frac{V_r s_\theta + V_s r_\theta}{V_r\sqrt{1-s_\theta^2} + V_s\sqrt{1-r_\theta^2}} \right). \tag{3.28}$$

In equations (3.27) and (3.28), $\nu := -\tan\phi$, with ϕ the reflector dip (measured clockwise positive with the horizontal), and $2h/(V_{P0}t_m)$ is the offset-to-depth ratio, with h the half-offset. Note that in 2D, we have $-\pi/2 < \theta_{s,r} < \pi/2$. Given a reflector dip-angle ϕ and an offset-to-depth ratio $2h/(V_{P0}t_m)$, the angles $\theta_{s,r}$ [and thus $(s,r)_\theta$] were found using the Gauss-Newton method (Dennis, 1977). Then, these values were used in the exact migration equations and in their weak-anisotropy approximations (with $p_{s,r} = (s,r)_\theta/V_{s,r}$) for $\eta = 0.0833$, $V_{NMO}(0) = 3.29$ km/s, $V_{P0} = 3$ km/s, and $f = \frac{5}{9}$. Figure 3.3 shows that indeed for all offset-to-depth ratios used, the approximations are highly accurate up to about 60 degrees for $\eta = 0.0833$, with the approximation to the migrated slope p_m losing accuracy fastest; i.e., for dips up to 60 degrees the approximations are accurate within 5% for offset-to-depth ratios up to 3, at least for η as large as 0.0833. Interestingly, for the migrated time t_m and horizontal displacement s_m, there is a combination of (nonzero) reflector dip and offset-to-depth ratio where the weak-anisotropy approximate expressions for t_m and x_m are exact.

In addition, Figure 3.4 shows contoured values of the exact equations normalized by their elliptic ($\eta = 0$) counterparts [i.e. (a) $t_m/(t_m|_{\eta=0})$, (b) $s_m/(s_m|_{\eta=0})$ and (c) $p_m/(p_m|_{\eta=0})$], as a function of reflector dip and offset-to-depth ratio, for the same values of η, $V_{NMO}(0)$, V_{P0}, f, and t_m used in Figure 3.3. Note that for zero offset-to-depth ratio Figure 3.4 is the same as the drawn black lines in Figure 3.2a. Clearly, ignoring the presence of anisotropy (or, better, anellipticity) leads to error in the migrated time, lateral location, and orientation of the reflector. Noticeably, this error increases with increasing offset-to-depth ratio and increasing reflector dip. Therefore, the previous comment regard-

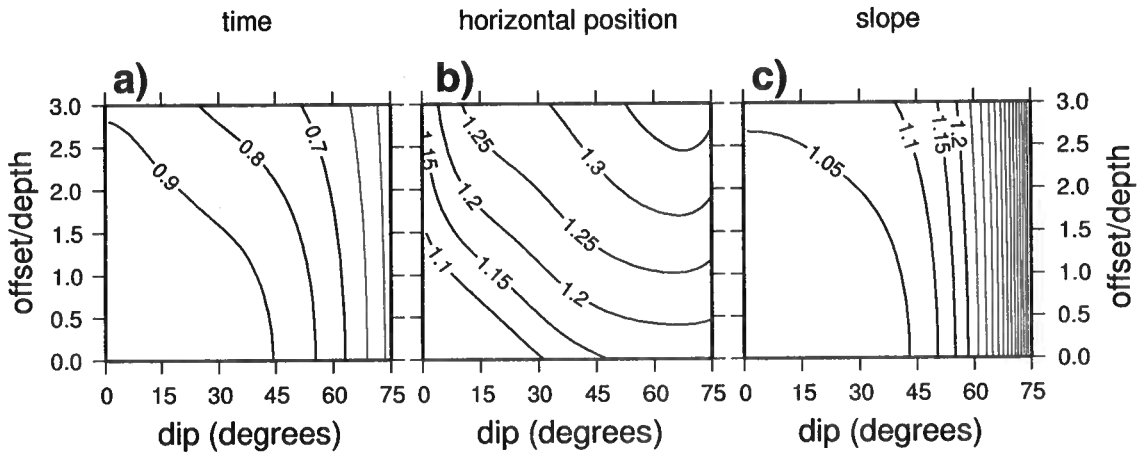


Figure 3.4. (a) $t_m/(t_m|_{\eta=0})$, (b) $s_m/(s_m|_{\eta=0})$ and (c) $p_m/(p_m|_{\eta=0})$ for 2D pre-stack map migration, as a function of reflector dip and offset-to-depth ratio. The values of the relevant parameters used to calculate this Figure are $\eta = 0.0833$, $t_m = 2$ s, with $V_{NMO}(0)$, V_{P0} and f the same as in Figure 3.1.

ing the importance of accounting for the presence of anisotropy with regard to positioning a borehole, that was made in the context of post-stack migration, becomes even more relevant in the context of pre-stack migration. Pre-stack migration error as a result of ignoring subsurface anisotropy has also been studied by Jaramillo & Larner (1995).

3.6 Pre-stack map demigration

While pre-stack map time-migration in homogeneous VTI media is a straightforward generalization of the post-stack problem, it is shown in Chapter 2 that the pre-stack map time-demigration problem involves the solution of a nonlinear set of four equations with four unknowns, i.e., two azimuth angles and two dips, that cannot be solved in closed-form. It turns out that even using the weak-anisotropy approximation does not allow for a solution in closed-form. While the system of nonlinear equations could be written out under the weak-anisotropy approximation, in the interest of brevity I refrain from such a treatment.

3.7 Conclusion and discussion

Since many rocks exhibit weak anisotropy only, I present weak-anisotropy approximations to the post-stack and pre-stack map time-migration equations previously developed by Douma & Calvert (2006) [see Chapter 2] for qP-waves in homogeneous VTI media. For the demigration case, the approximations are given only for zero-offset. The resulting expressions depend explicitly on η and $V_{NMO}(0)$ only. This confirms the previously reported

main dependence of the kinematics of qP-waves in (laterally homogeneous) VTI media on η and $V_{NMO}(0)$, but does so in the context of map time-migration and thus in the context of time-migration velocity analysis. This main dependence is also confirmed numerically for the exact expressions, which puts it in the context of map time-migration. At least for η values up to 0.08, the errors of the weak-anisotropy approximations are within 5%, for reflector dip up to 60 degrees and offset-to-depth ratio up to three. In addition, the numerical examples indicate that, even in the presence of weak anisotropy only, the errors in time imaging resulting from ignoring the presence of anisotropy are substantial. This indicates that accurately accounting for anisotropy seems highly relevant for accurate positioning of boreholes.

The presented equations allow for inversion of the anellipticity parameter η in the context of time-migration velocity analysis. By using data from two (or more) different offsets, we can use the difference between the migrated times, lateral positions and slopes (i.e., all used simultaneously), as an objective function to be minimized, by changing η. The optimum value for η then results in the minimum value of objective function. Such a scheme would be similar to the approach of Iversen *et al.* (2000), who use qP and qSV data to invert for the Thomsen anisotropy parameters ϵ and δ from map migration via ray tracing in heterogeneous anisotropic media. Using the expressions here, qP data could be used to invert for η in a time-migration setting, without significant computational effort. The resulting estimates of η could then be used as initial estimates in a more advanced inversion scheme such as the one presented by Iversen *et al.* (2000). This would help constrain the latter inversion.

Conventionally, estimation of η is done based on the deviation of the observed moveout in common-midpoint (CMP) gathers from being hyperbolic, i.e., based on velocity analysis in CMP gathers. Migration velocity analysis allows the estimated velocity to be related to a subsurface location, rather than a midpoint location. Because the presented equations allow the estimation of η to be done in a migration-analysis context, they can be expected to provide improved estimates of η in the subsurface when compared to moveout analysis in CMP gathers. Also, the combined estimation of migrated time, location, and orientation is expected to be more sensitive to changes in η than is moveout in CMP gathers. Hence, estimation of η based on map migration can be expected to provide better resolution in η than the resolution obtained from moveout analysis in CMP gathers.

Chapter 4

A hybrid formulation of map migration and wave-equation-based migration using curvelets[1]

4.1 Summary

Recently, curvelets have been introduced in the field of applied harmonic analysis and shown to optimally sparsify smooth functions away from singularities along smooth curves. In addition, it has been shown that curvelets allow a sparse representation of wave propagators. Since the wavefronts in seismic data lie mainly along smooth surfaces (or curves in two dimensions) and since the imaging operator belongs to the class of operators that is sparsified by curvelets, curvelets are plausible candidates for simultaneous sparse representation of both the seismic data and the imaging operator. Here, the first curvelet-based seismic imaging method is presented, using common-offset time-migration as an example. It is shown that with curvelets the leading-order approximation (in angular frequency, horizontal wavenumber, and migrated location) to such migration becomes a simple transformation of the coordinates of the curvelets in the data, combined with amplitude scaling. This transformation is calculated using common-offset (CO) map time-migration, which uses the local slopes provided by the curvelet decomposition of the data. The accuracy of the method is verified using numerical examples that indicate that using the leading-order approximation only, provides a good approximation to CO time-migration. Even though the current work treats constant media only, the essence of the presented leading-order derivation applies to pre-stack depth-migration also. The combination of map migration with curvelets, as presented in this work, unifies the finite-frequency character of curvelets with the ability of map migration to handle caustics. Because wave-equation-based migration is not troubled by the presence of caustics and honors the finite-frequency character of the data, this work can be viewed as a hybrid formulation of map migration and wave-equation-based migration.

4.2 Introduction

In the high-frequency approximation, seismic waves propagate along rays in the subsurface. The local slopes of reflections in seismic data measured at the surface determine

[1]This chapter together with Appendix H will be submitted for publication to *Geophysics* as: Huub Douma and Maarten V. de Hoop, 2006, A hybrid formulation of map migration and wave-equation-based migration using curvelets.

(together with the velocity of the medium at the surface), the directions in which we need to look into the earth from the surface to find the location and orientation of the reflector in the subsurface where the reflectionw occurred. The determination of a reflector position *and* orientation from the location of a reflection in the data *and* the local slope, is generally referred to as map migration (Kleyn, 1977). For an overview of the literature on this topic, and for an explanation of the applicability condition of map migration, we refer to Douma & de Hoop (2006) (see Chapter 2).

Given the slopes at the source and at the receiver locations, map migration provides a one-to-one mapping from the surface seismic measurements, i.e., locations, times, and slopes, to the reflector position and orientation in the image (provided the medium does not allow different reflectors to have identical surface seismic measurements that persist in being identical under small perturbations of the velocity model; see Douma & de Hoop (2006) for an explanation of this condition). This is in sharp contrast to migration techniques that do not make explicit use of the slopes in the data, such as Kirchhoff methods, where the data are summed over diffraction surfaces [see, e.g., Bleistein *et al.* (2000)]; such mappings are many-to-one in a way that all points along the diffraction surface are mapped to one diffractor location. The benefit of the explicit use of the local slopes in the data, is exploited in several seismic applications such as parsimonious migration (Hua & McMechan, 2001, 2003), controlled directional reception (CDR) (Zavalishin, 1981; Harlan & Burridge, 1983; Sword, 1987; Riabinkin, 1991), and stereo tomography (Billette & Lambaré, 1998; Billette *et al.*, 2003). This list is not complete and many more applications exist. In all these methods, the slopes are estimated from the data using additional processing techniques such as local slant-stacking, multidimensional prediction-error filters (Claerbout, 1992, p.186-201) or plane-wave destruction filters (Fomel, 2002; Claerbout, 1992, p.93-97).

Recently, in the field of applied harmonic analysis (Candès & Guo, 2002; Candès & Donoho., 2000; Candès & Donoho, 2004b) a tight frame of curvelets has been introduced (see Appendix I for an explanation of tight frames) which provides an essentially optimal representation of objects that are twice continuously differentiable (C^2) away from discontinuities along C^2 edges. Due to the wave character of seismic data, the reflections recorded in seismic data lie predominantly along smooth surfaces (or curves in 2D), just as geologic interfaces in the subsurface lie primarily along smooth surfaces. Therefore, it is plausible to assume that seismic data and their images can be sparsely represented using curvelets. This was earlier also noted by Herrmann (2003a,b). At points where the recorded wavefronts or the subsurface contain point-like discontinuities (e.g., at the edges of a fault in the subsurface), however, the level of sparsity that can be achieved with a curvelet representation naturally will be somewhat less than the sparsity achieved for the smooth parts of wavefronts or geologic interfaces. Because curvelets are anisotropic inseparable 2D extensions of wavelets that have main associated local directions, i.e., oriented wavepackets, using curvelets as building blocks of seismic data the local slopes in the data are built into the representation of the data; a simple projection of the data onto the curvelet frame (combined with an intelligent thresholding scheme to separate signal from noise) then provides the local directions associated with the recorded wavefronts.

Smith (1998) and Candès & Guo (2002) have shown that curvelets sparsify a class of

Fourier integral operators from which the seismic imaging operator can be constructed. In combination with the observation that reflections in seismic data lie mainly along smooth curves, it seems therefore that curvelets are plausible candidates for simultaneous compression of seismic data and the imaging operator. Waves with a given dominant wavelength are sensitive to variations in the medium with certain lengths scales only (Jannane *et al.*, 1989); i.e., a wave with a dominant wavelength of say 100 meters is hardly sensitive to variations in the medium on the scale of one centimeter (because the first Fresnel zone of a wave is proportional to the square-root of the wavelength). Just like wavelets, curvelets have a bandlimited character and thus have an associated dominant frequency. Because of this bandlimited nature, curvelets with different dominant frequencies are sensitive to variations in the medium at different scales. This allows the possibility to smooth the background velocity with filters related to the dominant wavelength of the curvelets (i.e., the scale of the curvelets), and propagate curvelets of different scales through different smoothed versions of the medium. Smith (1998) has shown that the propagation of a curvelet through such a smoothed medium is governed by the Hamiltonian flow in this smoothed medium associated with the center of the curvelet. Here the center of a curvelet is defined by its center in phase-space, meaning the central location of a curvelet combined with its central direction. This means that a curvelet is treated as if it was a particle with an associated momentum (or direction). For each smoothed medium, this observation reduces to the statement that the propagation of a curvelet (through an infinitely smooth medium) is well-approximated by translating the center of the curvelet along the corresponding Hamiltonian flow, i.e., translating it along the ray associated with the center of a curvelet (Smith, 1998; Candès & Demanet, 2005). In fact, the procedure just outlined yields a leading-order contribution to the solution of the wave equation (Smith, 1998). Hence this procedure admits wave-equation-based seismic imaging with curvelets.

In this work we focus on the simple case of homogeneous media to learn the basic characteristics of seismic imaging with curvelets and as preparation for imaging in heterogeneous media with curvelets. Because in homogeneous media the above-mentioned smoothing is unnecessary, we ignore such smoothing in this paper altogether. Instead, we focus our attention on verifying, in the context of seismic imaging, the statement that curvelets can be treated as particles with associated directions (or momenta). We show that by treating curvelets as such, common-offset (CO) time-migration becomes, to leading order in angular frequency, horizontal wavenumber, and migrated location, a transformation of the coordinates (i.e., midpoint and time) of each curvelet in the data, combined with amplitude scaling. The coordinate transformation can be calculated with the aid of the explicit formulas for CO map time-migration developed by Douma & de Hoop (2006) (see Chapter 2). Numerical verification of the proposed leading-order approximation based on this simple coordinate transformation confirms that it provides a good approximation to CO time-migration. We emphasize that the essence of the derivation is not limited to homogeneous media but mention that the leading-order approximation to pre-stack depth-migration is obtained also by a transformation of coordinates. In that case, however, the accuracy of the leading-order approximation is currently unknown.

The coordinate transformation presented below is calculated using map migration. Be-

cause map migration makes explicit use of the slopes in the data, it constitutes a mapping in phase-space (Douma & de Hoop, 2006). Moreover, since there are no caustics in phase-space, map migration naturally handles caustics. The combination of map migration with curvelets, as presented in this work, unifies the ability of map migration to handle caustics with the finite frequency character of curvelets. The ability to handle caustics while honouring the finite-frequency character of the data, is also found in wave-equation-based migration [e.g., Claerbout (1970, 1976), Stoffa *et al.* (1990), Biondi & Palacharla (1996), Le Rousseau & de Hoop (2001)]. Therefore, we view this work as a hybrid formulation of map migration and wave-equation-based migration.

The outline of this paper is as follows. First we explain what curvelets are, how they are constructed, and what are their main properties. Subsequently we present examples of digital curvelets from the digital curvelet transform (Candès *et al.*, 2005). A more detailed treatment of a particular construction of real-valued curvelets is included in Appendix H. Then, we show an example of the use of curvelets as building blocks of seismic data, and explain the relationship between curvelets and map migration. We proceed to derive that to leading order in angular frequency, horizontal wavenumber, and migrated location, CO time migration is equivalent to a transformation of the coordinates of the curvelets in the data combined with amplitude scaling, and analyze the components that constitute the coordinate transformation. Subsequently we explain how this transformation can be calculated from pre-stack map migration. Finally, we verify the accuracy of the proposed method using numerical examples, and finish with a discussion of the results.

4.3 Curvelets

In this section we explain what curvelets are, how they are constructed, and what are their main properties. Appendix H, which provides a more detailed treatment of a particular construction of curvelets, closely follows the original treatment of the construction of real-valued curvelets by Candès & Donoho (2004b) but provides additional explanations and derivations to guide the non-specialist. For a short summary of the more general complex-valued curvelets, we refer the reader to Candès & Demanet (2005), while Candès *et al.* (2005) describe two different implementations of digital curvelet transforms. A frame similar to the curvelet frame was earlier introduced by Smith (1998).

In wavelet theory [e.g., Mallat (1998)], a 1D signal is decomposed into wavelets, where a wavelet is localized in both the independent variable *and* its Fourier dual, say time and frequency. Such localization is understood within the limits imposed by the Heisenberg uncertainty principle. These wavelets can be translated along the time axis through a translation index, and dilated in their frequency content through a scale index. They are uniquely determined by both indices: the translation index m determines their location along, say, the time axis, while the scale index j determines their location along, say, the frequency axis.

Curvelets are basically 2D *anisotropic* (see below) extensions to wavelets that have a main direction associated with them. Analogous to wavelets, curvelets can be translated and dilated. The dilation is given also by a scale index j that controls the frequency content

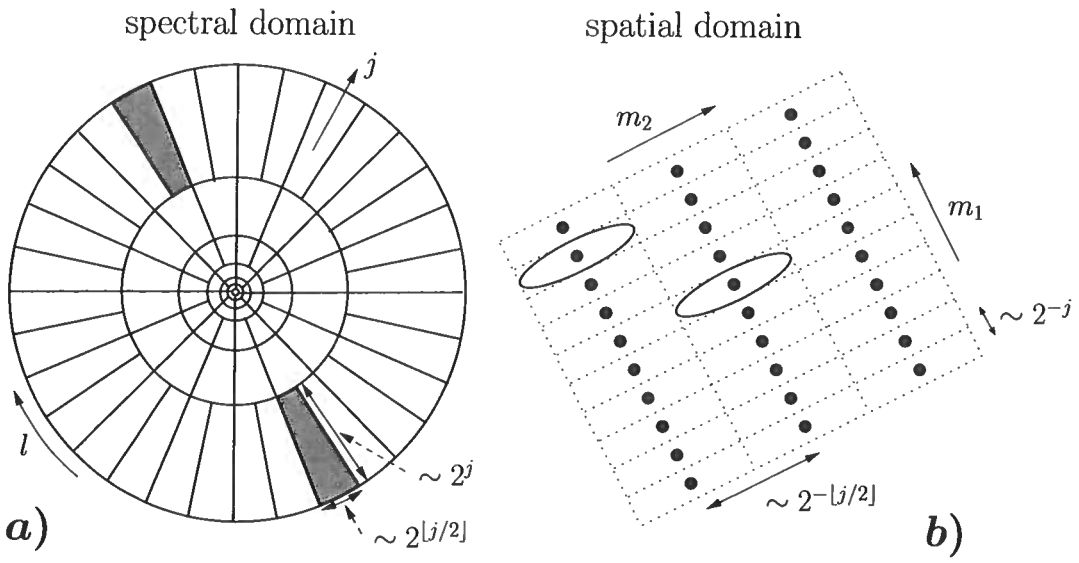


Figure 4.1. Tilings of the curvelet frame in the spectral domain (a) and the spatial domain (b). In the frequency domain a curvelet is supported near a wedge on a polar grid (i.e., the actual support extends slightly further than the indicated wedge), where the width of the wedge is proportional to $2^{\lfloor j/2 \rfloor}$ and its length is proportional to 2^j. On the support of such a wedge, a local Fourier basis provides a Cartesian tiling of the spatial domain (shown schematically in b). The essential support of a curvelet in the spatial domain is indicated by an ellipse (while again the actual support extends beyond this ellipse).

of the curvelet, while the translation is indexed by m_1 and m_2 (in two dimensions)[2]. A curvelet has a main associated direction that can be changed through a rotation. This rotation is indexed by an angular index l. The relation between these indices and the location of the curvelet in the spatial and spectral domains is shown in Figures 4.1a and b. A curvelet is uniquely determined by all four indices (j, l, m_1, m_2).

As explained in Appendix H, curvelets satisfy the anisotropic scaling relation width $\propto$ squared length in the spatial domain[3]. This is generally referred to as *parabolic scaling.* This anisotropic scaling relation associated with curvelets is a key ingredient to the proof in Candès & Donoho (2004b) that curvelets provide the sparsest representations of C^2 (i.e., twice continuously differentiable) functions away from edges along piecewise C^2 curves. The search for sparse representations of such functions in the field of image analysis was the original motivation for their construction, as wavelets fail to sparsely represent such functions (Candès & Donoho, 2004b) due to their isotropic character. The anisotropic scaling relation is a key difference between wavelets and curvelets. The parabolic scaling is also a key ingredient to the proof that curvelets remain localized in phase-space (i.e., remain curvelet-like) under the action of the wave operator provided the medium is smoothed appropriately prior to propagation (Smith, 1998).

Curvelets are constructed through the following sequence of operations. First, the spectral domain is bandpass filtered in the radial direction into dyadic annuli (or subbands); this means that the radial widths of two neighboring annuli differs by a factor of two, the next outer annulus having twice the radial width of the inner one (see Figure 4.1a). Each subband is subsequently subdivided into angular wedges, where the number of wedges in each subband is determined by the frequency content (or the scale index j) of the subband (see Figure 4.1a). The number of wedges in a subband with scale j is $2^{\lfloor j/2 \rfloor}$, where the notation $\lfloor p \rfloor$ denotes the integer part of p. This means that the number of wedges in a subband increases only every other scale[4]. Subsequently, each wedge is multiplied by a 2D orthonormal Fourier basis for the rectangle that just covers the support of the wedge. According to the discrete Fourier transform, this basis has the fewest members if the area of this rectangle is minimum, since then the product of both sampling intervals in space is largest. Therefore, the orientation of this rectangle rotates with the angular wedge, and the spatial tiling associated with the local Fourier basis is oriented along the central direction of the angular wedge (see Figure 4.1b); that is, the spatial tiling associated with each angular wedge depends on the particular orientation of the wedge. The subband filtering

[2]We differ from the standard notation k_1 and k_2 to avoid confusion with the wave-vector components.

[3]Considering the difference in dimension between width and squared length, this scaling relation holds when both width and length are made dimensionless through division by a reference length, typically the sampling interval in numerical implementations. Otherwise this relation would involve a proportionality constant that has a dimension which adjusts for the difference in dimension between width and squared length.

[4]This is a consequence of the dyadic nature of the subband filtering done in the first step combined with the desired parabolic scaling. Without this splitting at every other scale resulting from the parabolic scaling, the number of wedges would increase with a factor $\sqrt{2}$ if the scale index was increased by one. In this case we would not have an integer number of wedges, hence the doubling of the number of angular wedges every other scale only.

gives curvelets their bandlimited nature (just as with wavelets), whereas the subdivision of these subbands into angular wedges provides them with orientation. The local Fourier basis over the support of the angular wedge allows the curvelets to be translated in space. Therefore, curvelets are in essence a tiling of phase-space; i.e., a tiling of two variables and their two Fourier duals. The tiling is non-trivial in that the sampling of phase space is polar in the spectral domain, but Cartesian in the spatial domain. As explained in Appendix H, curvelets are essentially Heisenberg cells in phase-space.

Roughly speaking, we can think of curvelets as small pieces of bandlimited plane waves. The difference between this rough description and the actual interpretation lies, of course, in the fact that a bandlimited plane wave has associated with it one $\boldsymbol{k}$ direction only, whereas a curvelet is associated with a small range of $\boldsymbol{k}$ vectors. A better description is the term *coherent wave packet* which was around before curvelets were ever constructed [e.g. Smith (1997, 1998)] and dates back to the work of Córdoba & Fefferman (1978). The frequency-domain tiling of the curvelet frame is the same as the dyadic parabolic decomposition or second dyadic decomposition (Gunther Uhlmann, personal communication with Maarten de Hoop) used in the study of Fourier integral operators [see e.g. Stein (1993)], which was around long before the construction of the curvelet frame (Fefferman, 1973).

Curvelets form a *tight frame* for functions in $L^2\left(\mathbb{R}^2\right)$ — see Appendix H for the proof of this property and Appendix I for an introduction to tight frames. This means that, much as with an orthonormal basis, we have a reconstruction formula

$$f = \sum_{\mu \in \mathcal{M}} (f, c_\mu)\, c_\mu \,, \quad (f, c_\mu) = \int_{\mathbb{R}^2} f(\boldsymbol{x}) c_\mu^*(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} \,, \tag{4.1}$$

where c_μ denotes a curvelet with multi-index $\mu = (j, l, m_1, m_2)$, the superscript $*$ denotes complex conjugation, $\mathcal{M}$ is a multi-index set, and $f(x_1, x_2) \in L^2\left(\mathbb{R}^2\right)$. Thus, we can express an arbitrary function in $L^2\left(\mathbb{R}^2\right)$ as a superposition of curvelets. The quantity (f, c_μ) is the coefficient of curvelet c_μ which denotes the projection of the function f on curvelet c_μ.

4.3.1 Digital curvelets

In the construction of curvelets treated so far, the sampling of the spectral domain is done in polar coordinates while the sampling of the spatial domain is Cartesian (see Figures 4.1a and b). From a computational point of view, this combination is not straightforward to implement. Combining Cartesian coordinates in both domains is straightforward and is standard in data processing. Therefore, for the purpose of digital curvelet transforms, the polar coordinates in the spectral domain are replaced with Cartesian coordinates (see Figures 4.2a). Also, in the field of image analysis [where the digital curvelet transform was originally developed (Candès & Donoho, 2004a; Candès *et al.*, 2005)], images usually have Cartesian spatial coordinates to begin with; hence it is natural to have Cartesian coordinates in the spectral domain also. This allows straightforward application of Fast Fourier Transform algorithms. This argument holds for seismic as well as image data.

To go from polar coordinates to Cartesian coordinates in the spectral domain, the concentric circles in Figure 4.1a are replaced with concentric squares (see Figure 4.2a);

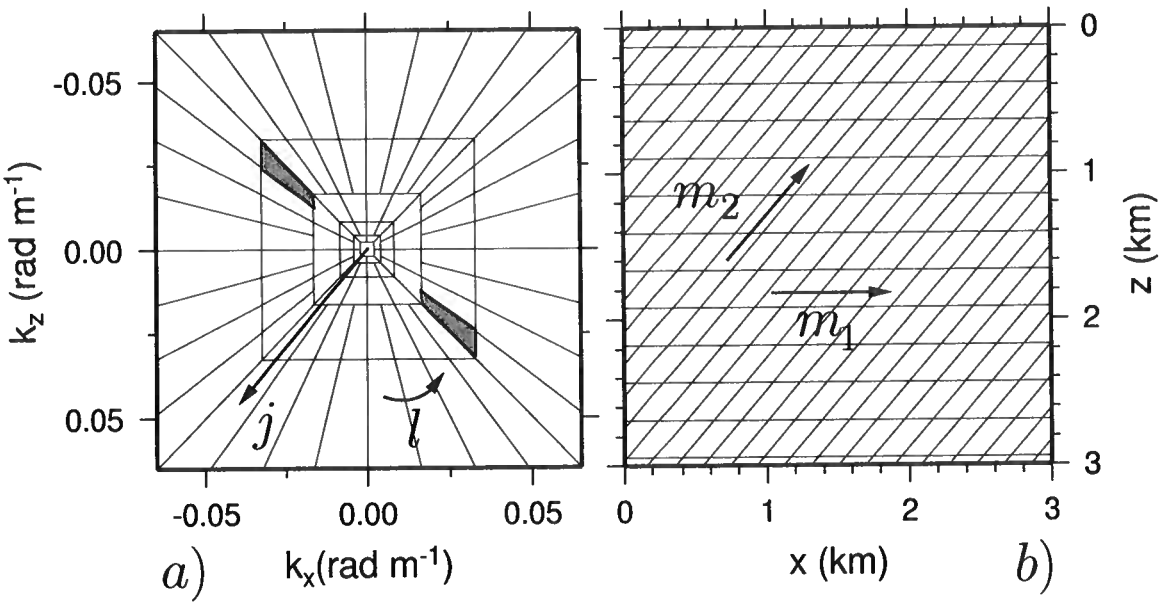


Figure 4.2. Tilings for digital curvelets in the spectral domain (a) and the spatial domain (b). For digital curvelets, the concentric circles in the spectral domain are replaced with concentric squares, and the Cartesian spatial grid is sheared.

hence the rotational symmetry is replaced with a sheared symmetry. As a consequence, the Cartesian sampling in the spatial domain is a Cartesian grid that is sheared rather than rotated (cf., Figures 4.1b and 4.2b)[5]. This construction allows a rapidly computable digital curvelet transform (Candès *et al.*, 2005). For more details on the implementation of digital curvelet transforms, we refer to Candès & Donoho (2004a) and Candès *et al.* (2005).

4.3.2 Examples of digital curvelets

Figure 4.3 shows examples of digital curvelets. The left column shows curvelets in the spatial domain, while the right column shows their associated spectra. Superimposed on the spectra is the spectral tiling of the digital curvelet transform. The middle column shows the associated spatial lattice for each of the curvelets, where the centers of the cells are the locations of the centers of the curvelets in space[6]. Here the spatial cells on the spatial lattice are colored according to the magnitude of the curvelet coefficient (here always unity); black equals one and white equals zero. Figure 4.3b shows a translated version of the curvelet in Figure 4.3a; the spectral tile is the same, but the spatial tile has changed, i.e., indices j and l are held constant, but the translation indices m_1 and m_2 are different. Figure 4.3c shows a curvelet with the same spatial location and the same scale index as that in Figure 4.3a, but with a different angular index l. That is, the spectral tile has moved within the same filter band, i.e., within the same concentric squares. Note how the spatial lattice changes as we change the angular index l. Finally, Figure 4.3d shows a dilated version of the curvelet shown in Figure 4.3a; the spatial location is the same, but the spectral tile has moved outward into a neighboring annulus (or subband), i.e., the scale index $j \rightarrow j+1$. Since the neighboring annulus is subdivided into more wedges, the angular index l has also changed, but in such a way that the direction of the curvelet is basically the same. Similarly, since the larger scale has a finer spatial sampling associated with it, the translation indices m_1 and m_2 have also changed, but in such a way that the curvelet location is the same.

4.4 Curvelets remain curvelet-like when subjected to the class of operators relevant for seismic imaging

The action of operators belonging to the class of Fourier integral operators that can be sparsely represented using curvelets, which includes the CO time-migration operator, can be described in terms of propagation of singularities along a Hamiltonian flow (Smith, 1998; Candès & Guo, 2002). The action on a curvelet of a particular scale can be approximated by translating the curvelet along the ray associated with the center of the curvelet through the medium smoothed for that particular scale. This means that, in the appropriately smoothed media, curvelets remain fairly localized in both the spatial domain and the spectral domain.

[5]Here the centers of the cells are the locations of the centers of the curvelets in space.

[6]Throughout this work, we used the nonequispaced fast Fourier transform-based curvelet transform (Candès *et al.*, 2005). Because this transform induces a sheared spatial grid, as opposed to the wrapping-based transform (Candès *et al.*, 2005), which induces a rectangular grid, all spatial lattices shown in this work are sheared grids.

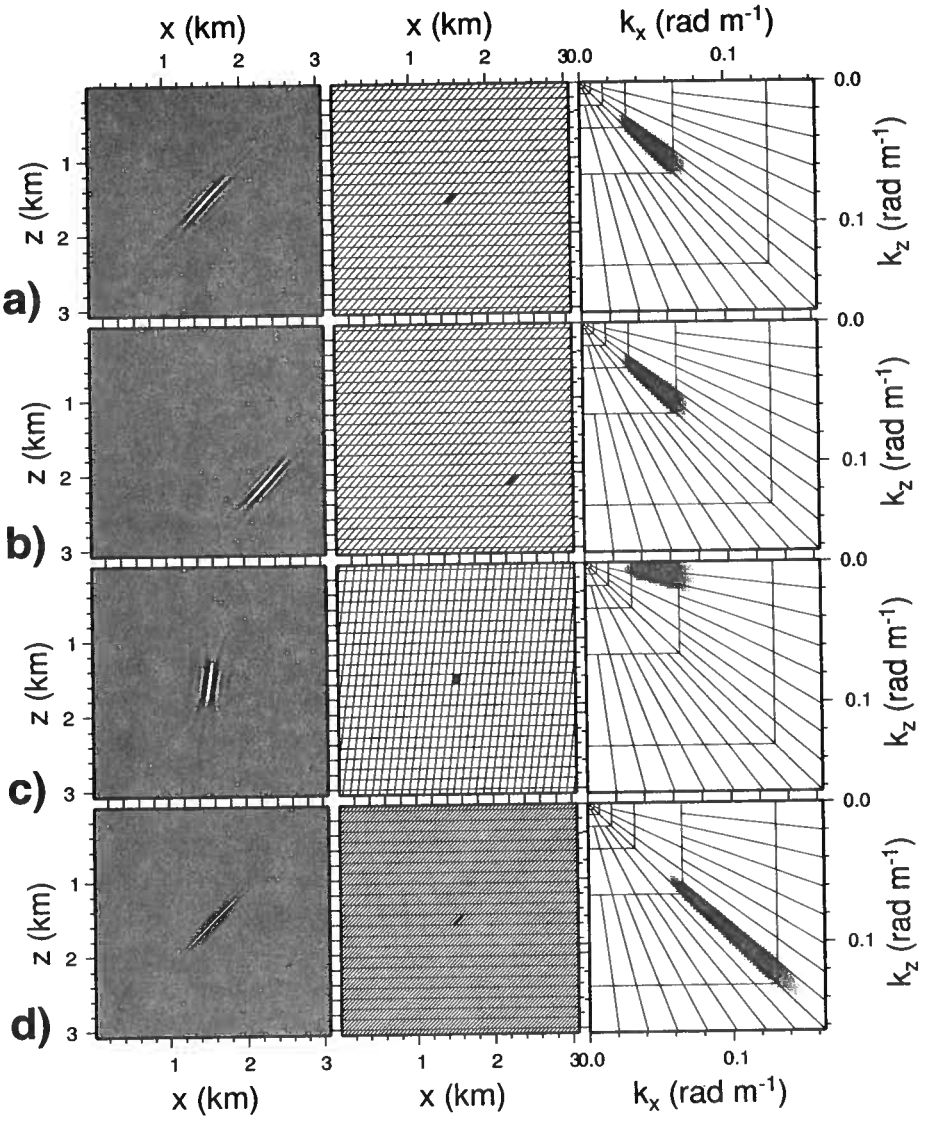


Figure 4.3. First column: curvelets in the spatial domain. Second column: associated spatial lattices, and spatial cells colored according to the value of the coefficient (black is one, white is zero). Third column: amplitude spectra and frequency-domain tilings. First row: a curvelet. Second row: curvelet from (a) with different translation indices. Third row: curvelet from (a) with a different angular index. Fourth row: curvelet from (a) with a different (higher) scale index (here the translation indices and the angular index are also different because they both depend on the scale index).

Hence, the propagated curvelet can be constructed by using neighboring curvelets only, where neighboring is understood in the context of phase-space; i.e., a neighboring curvelet is a curvelet that is close in the spatial domain and has orientation close to the orientation of the curvelet that is propagated along the central ray, i.e., the ray associated with the center of the curvelet in phase-space. Note that propagation of a curvelet in accordance with the Hamiltoninan is an operation in phase-space since it makes explicit use of the slope of the curvelet. Since in phase-space no caustics can occur, curvelet-based imaging does not even notice the formation of caustics.

For homogeneous media smoothing is unnecessary and curvelets of all scales can be propagated along the central ray through the same medium. To illustrate this, Figure 4.4 shows the result of CO Kirchhoff migration of a curvelet [taken from Douma & de Hoop (2004)]. The top row shows the input curvelet in space (the vertical axis was converted to depth using $z = vt_u/2$ for convenience, with z denoting depth, and v and t_u denoting velocity and the two-way traveltime, respectively) and its associated amplitude spectrum. Again the spatial distribution of the coefficients is shown in the middle panel, just as in Figure 4.3. The left-most panel of the bottom row shows the CO Kirchhoff-migrated curvelet. Notice how the migrated curvelet is clearly localized in space and determines *only part of the isochron*, in sharp contrast to the whole isochron if a single sample (or a 'spike') would be used as input to the migration. This confirms that in the context of migration (at least for CO time-migration), curvelets are indeed a more appropriate choice for building blocks of seismic data than are spikes (that are currently used to represent seismic data). The spectrum of the migrated curvelet (bottom right) is clearly also localized after the migration, and overlies only four wedges in the curvelet tiling of the spectrum. The middle panel shows the coefficients for the spatial area in the lower left quadrant of the leftmost figure (outlined by the dotted lines), for the wedges labeled '1' through '4'; since curvelets of different orientations induce different spatial tilings[7], and since the migrated curvelet overlies four wedges, the spatial distribution for the coefficients related to all four wedges are shown separately in the lower middle pannel of Figure 4.4. Indeed this reveals that the migrated curvelet is constructed from several curvelets, but only few of them having large coefficients. This confirms that curvelets remain fairly localized in both the spatial and spectral domain (i.e., they remain curvelet-like), at least after CO time-migration.

4.5 Curvelets as building blocks of seismic data

Seismic reflections in seismic data lie primarily along smooth surfaces (or curves in two dimensions). Even diffractions from discontinuities in the earth's subsurface, such as edges of geologic interfaces caused by faulting, lie along smooth surfaces. This is a direct consequence of the wave-character of seismic data. As mentioned in the introduction, it is intuitive that curvelets can be used to sparsely represent seismic data because curvelets provide the sparsest representations of smooth (C^2) functions away from edges along piecewise C^2 curves

[7]The wrapping-based curvelet transform (Candès *et al.*, 2005) does not induce different tilings for spectral wedges with the same scale index but different angular indices.

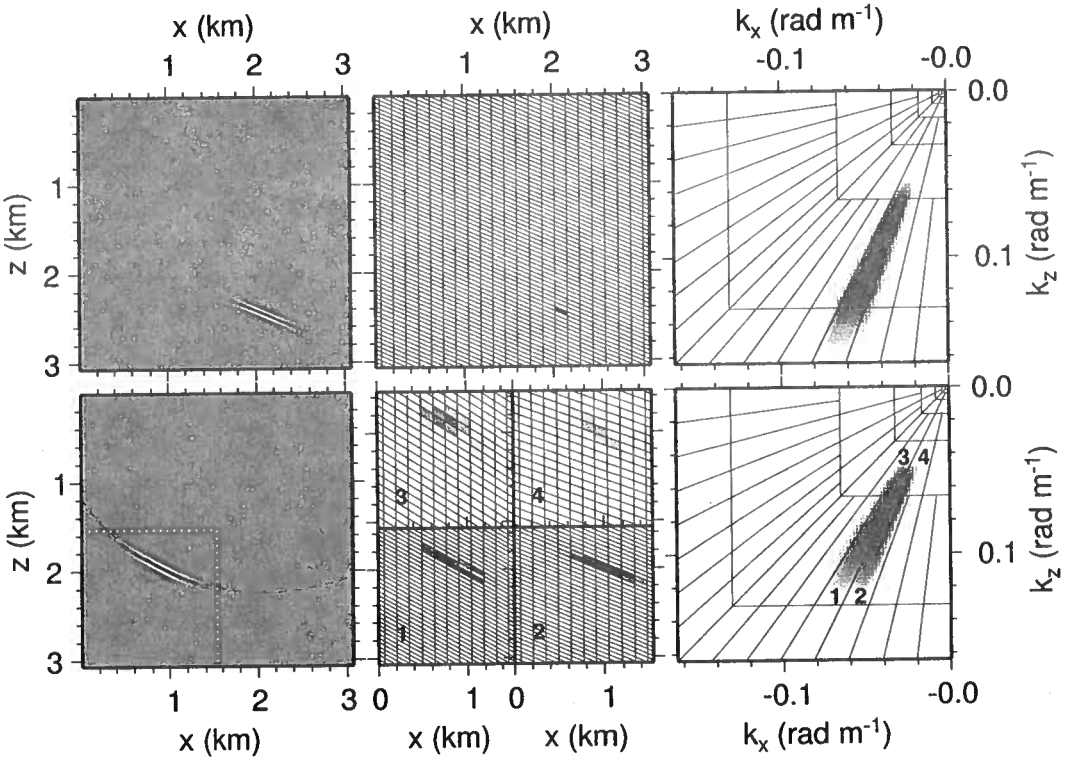


Figure 4.4. Top row: a curvelet with a dominant frequency of about 30 Hz (left, shown in depth $z = vt_u/2$, for consistency), the normalized absolute value of the coefficient on the spatial lattice (middle), and its amplitude spectrum (right). Bottom row: CO Kirchhoff migration of the curvelet in the top row. The middle panel in this row shows the coefficients on the spatial lattice in the lower left quadrant of the leftmost panel (indicated with the dotted lines in the leftmost panel) for each of the numbered wedges (labeled '1' to '4') in the spectrum (right).

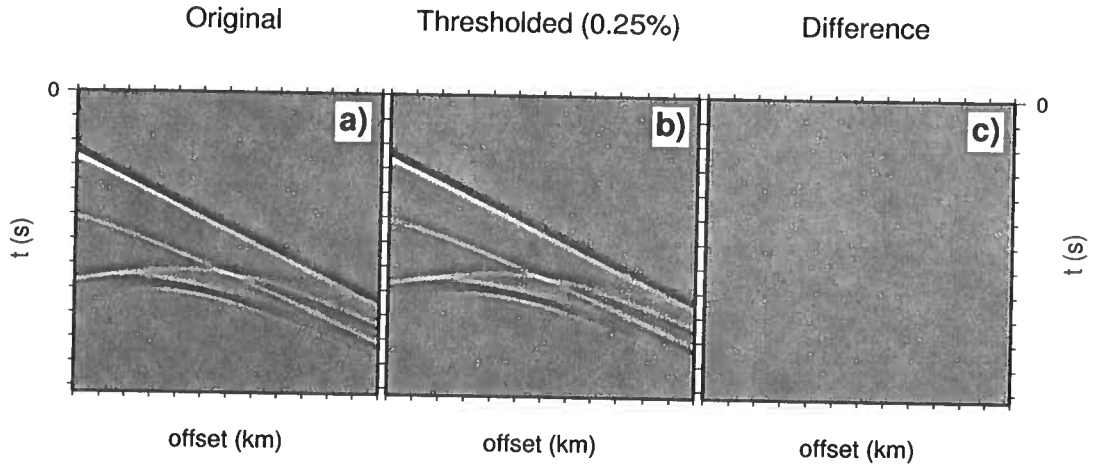


Figure 4.5. Synthetic common-shot gather with cusped wavefront: original (a), reconstructed using only the 0.25% largest curvelet coefficients (b), and the difference (c). The reconstruction with 0.25% of the curvelets is almost identical to the original common-shot gather. In this example, using only 0.25% of the curvelets results in about 30 times fewer curvelets than input samples in the gather.

(Candès & Donoho, 2004b). Throughout this work, we adopt this intuition and illustrate this with a simple synthetic example. Hence, we focus on the sparse representation of the imaging operator rather than the data.

Figure 4.5a shows part of a synthetic common-shot gather, where the wavefront has a cusp. These data relate to a model with a syncline-shaped reflector. Figure 4.5b shows the reconstructed gather where only the 0.25% largest curvelet coefficients were used. For the particular example shown, this relates to a compression ratio of about 30; i.e., we used 30 times fewer curvelets than the number of sample values in the original gather to reconstruct the data. From Figure 4.5c the difference between the original and reconstructed data is negligible. Note that this large compression ratio is partly a result of the synthetic data having many zero sample values to start with; for a field data example this compression ratio would likely be smaller. This example, however, shows that with curvelets as building blocks of seismic data, the data can be represented with fewer curvelets than samples in the data, and with essentially no residual, even in areas where the wavefront has cusps.

In this example, we have applied a hard thresholding to the data; we discarded 99.75% of the curvelets. At first sight one might therefore think that the compression ratio should be 400. The curvelet transform, however, is a redundant transform, meaning that if all curvelets are used to reconstruct the data, there are more curvelets than sample points in the data. Different digital implementations of the curvelet transform have different redundancies[8]

[8]In this particular example, the apparent compression ratio (400) and the associated implied redundancy

The quantity between the square brackets is an oscillatory integral that can be approximated using the method of stationary phase [e.g., Bleistein *et al.* (2000, p.129)]. Treating $-\omega$ as the formal large parameter and defining the phase Φ as

$$\Phi(\boldsymbol{y},x,k_x,\omega;h) := \frac{r_s(\boldsymbol{y},x;h)+r_g(\boldsymbol{y},x;h)}{v} - \frac{k_x}{\omega}x\,, \tag{4.11}$$

we get

$$\begin{aligned}\beta(\boldsymbol{y}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} &\frac{a(\boldsymbol{y},x(\boldsymbol{y},k_x/\omega;h);h)}{\sqrt{2\pi|\Phi''(\boldsymbol{y},x(\boldsymbol{y},k_x/\omega;h),k_x,\omega;h)|}} e^{-i\omega\Phi(\boldsymbol{y},x(\boldsymbol{y},k_x/\omega;h),k_x,\omega;h)} \\ &\times\hat{U}_s(k_x,\omega)e^{i(\pi/4)\,\mathrm{sgn}\,\omega[1-\mathrm{sgn}(\Phi''(\boldsymbol{y},x(\boldsymbol{y},k_x/\omega;h),k_x,\omega;h))]}\mathrm{d}k_x\mathrm{d}\omega\,,\end{aligned} \tag{4.12}$$

with $\Phi''(\boldsymbol{y},x(\boldsymbol{y},k_x/\omega;h),k_x,\omega;h) := \left.\dfrac{\partial^2\Phi}{\partial x^2}\right|_{x(\boldsymbol{y},k_x/\omega;h)}$ the Hessian evaluated at the stationary midpoint location $x(\boldsymbol{y},k_x/\omega;h)$. Note that the factor $\sqrt{|\omega|}$ in equation (4.10) has canceled out through the stationary phase approximation. Calculating the Hessian explicitly gives

$$\Phi'' = \frac{y_2^2}{v}\left(\frac{1}{r_s^3(\boldsymbol{y},x;h)}+\frac{1}{r_g^3(\boldsymbol{y},x;h)}\right)\,. \tag{4.13}$$

Since $r_s(\boldsymbol{y},x;h)$ and $r_g(\boldsymbol{y},x;h)$ represent distances (that are always positive) and because we have $y_2 \geq 0$, it follows that the Hessian is positive definite, whence $\mathrm{sgn}(\Phi'') = 1$. Using this in equation (4.12) gives

$$\begin{aligned}\beta(\boldsymbol{y}) \;=\; \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} &\frac{a(\boldsymbol{y},x(\boldsymbol{y},k_x/\omega;h);h)}{\sqrt{2\pi|\Phi''(\boldsymbol{y},x(\boldsymbol{y},k_x/\omega;h),k_x,\omega;h)|}} e^{-i\omega\Phi(\boldsymbol{y},x(\boldsymbol{y},k_x/\omega;h),k_x,\omega;h)} \\ &\times\hat{U}_s(k_x,\omega)\mathrm{d}k_x\mathrm{d}\omega\,.\end{aligned} \tag{4.14}$$

The stationary midpoint location $x(\boldsymbol{y},k_x/\omega;h)$ satisfies

$$\left.\frac{\partial\Phi(\boldsymbol{y},x,k_x,\omega;h)}{\partial x}\right|_{x=x(\boldsymbol{y},k_x/\omega;h)} = 0\,. \tag{4.15}$$

This condition leads to a quartic equation in x with $x(\boldsymbol{y},k_x/\omega;h)$ as its solution. This solution gives the midpoint location x as a function of the slope $p = k_x/\omega$ in the CO data and the migrated location $\boldsymbol{y}$. The resulting expression for $x(\boldsymbol{y},k_x/\omega;h)$ is thus a map demigration as a function of unmigrated (or input) as well as migrated (output) quantities; i.e., the midpoint location is determined as a function of the unmigrated slope and the migrated location. Note that to find the stationary midpoint location $x(\boldsymbol{y},k_x/\omega;h)$ the quartic equation does not need to be solved if we decompose the data into curvelets. In that case we can use the resulting central midpoint location (x), central traveltime (t) and central slope (p) to determine the migrated location ($\boldsymbol{y}$) through the map migration equations (4.2)–(4.4). In this way, for each input curvelet, we know the relation between

the midpoint location x on one side, and the migrated location $\boldsymbol{y}$ and the slope (p) in the data on the other side. Hence, we need not solve the quartic equation if we setup the migration in the data-domain (i.e., an input-based algorithm, see below) rather than in the image-domain (i.e., an output-based algorithm). The explicit form of the quartic equation is given in Appendix J.

Let $\beta_\mu(\boldsymbol{y})$ be the image of one input curvelet c_μ, with multi-index $\mu = (j, l, m_1, m_2)$. That is, we replace the input data $\hat{U}_s(k_x, \omega)$ in equation (4.14) with one curvelet $\hat{c}_\mu(k_x, \omega)$, where the notation $\hat{c}_\mu$ denotes the Fourier transform of the curvelet $c_\mu(x, t)$ in the $x - t$ domain. Therefore we have

$$\beta_\mu(\boldsymbol{y}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} a'(\boldsymbol{y}, k_x, \omega; h) e^{-iP(\boldsymbol{y}, x(\boldsymbol{y}, k_x/\omega; h), k_x, \omega; h)} \hat{c}_\mu(k_x, \omega) \mathrm{d}k_x \mathrm{d}\omega \,, \tag{4.16}$$

where we have defined

$$a'(\boldsymbol{y}, k_x, \omega; h) := \frac{a(\boldsymbol{y}, x(\boldsymbol{y}, k_x/\omega; h); h)}{\sqrt{2\pi |\Phi''(\boldsymbol{y}, x(\boldsymbol{y}, k_x/\omega; h), k_x, \omega; h)|}} \,, \tag{4.17}$$

and

$$\begin{aligned} P(\boldsymbol{y}, k_x, \omega; h) &:= \omega\, \Phi(\boldsymbol{y}, x(\boldsymbol{y}, k_x/\omega; h), k_x, \omega; h) \\ &= \omega\, t(\boldsymbol{y}, x(\boldsymbol{y}, k_x/\omega; h); h) - k_x\, x(\boldsymbol{y}, k_x/\omega; h) \,, \end{aligned} \tag{4.18}$$

with

$$t(\boldsymbol{y}, x; h) := (r_s(\boldsymbol{y}, x; h) + r_g(\boldsymbol{y}, x; h))/v \tag{4.19}$$

the traveltime from the source to the reflector and back to the receiver. Equation (4.16) is reminiscent of plane-wave migration (Akbar *et al.*, 1996).

Noticing that $\hat{c}_\mu(k_x, \omega)$ is localized near a wedge with center angular frequency ω_u and center horizontal wavenumber k_x^u, we linearize the dependence of P on ω and k_x around ω_u and k_x^u. This means that, apart from the amplitude a', the oscillatory integral on the right-hand side of equation (4.16) becomes simply an inverse Fourier transform. Doing this, we have

$$\begin{aligned} P(\boldsymbol{y}, k_x, \omega; h) \approx\; & P(\boldsymbol{y}, k_x^u, \omega_u; h) \\ & + (\omega - \omega_u) \left.\frac{dP}{d\omega}\right|_{(k_x^u, \omega_u)} + (k_x - k_x^u) \left.\frac{dP}{dk_x}\right|_{(k_x^u, \omega_u)} \,, \end{aligned} \tag{4.20}$$

where the derivatives $dP/d\omega$ and dP/dk_x denote total derivatives. From equation (4.18) it is immediate that

$$P(\boldsymbol{y}, k_x^u, \omega_u; h) = \omega_u\, t(\boldsymbol{y}, x(\boldsymbol{y}, p_u; h); h) - k_x^u\, x(\boldsymbol{y}, p_u; h) \,, \tag{4.21}$$

where $p_u := k_x^u/\omega_u$. Then, calculating the total derivative, it follows from equation (4.18)

that

$$\left.\frac{dP}{d\omega}\right|_{(k_x^u,\omega_u)} = t(\boldsymbol{y}, x(\boldsymbol{y}, p_u; h); h) + \left.\left(\omega \frac{\partial t}{\partial x}\frac{\partial x}{\partial \omega}\right)\right|_{(k_x^u,\omega_u)} - \left.\left(k_x \frac{\partial x}{\partial \omega}\right)\right|_{(k_x^u,\omega_u)} . \tag{4.22}$$

Since $p = \partial t/\partial x$ and that $\omega p = k_x$, it follows that

$$\left.\frac{dP}{d\omega}\right|_{(k_x^u,\omega_u)} = t(\boldsymbol{y}, x(\boldsymbol{y}, p_u; h); h) . \tag{4.23}$$

Similarly, it follows that

$$\begin{aligned} \left.\frac{dP}{dk_x}\right|_{(k_x^u,\omega_u)} &= \left.\left(\omega \frac{\partial t}{\partial x}\frac{\partial x}{\partial k_x}\right)\right|_{(k_x^u,\omega_u)} - x(\boldsymbol{y}, p_u; h) - \left.\left(k_x \frac{\partial x}{\partial k_x}\right)\right|_{(k_x^u,\omega_u)} \\ &= -x(\boldsymbol{y}, p_u; h) . \end{aligned} \tag{4.24}$$

Using equations (4.21), (4.23), and (4.24) in (4.20), we have

$$P(\boldsymbol{y}, k_x, \omega; h) \approx \omega\, t(\boldsymbol{y}, x(\boldsymbol{y}, p_u; h); h) - k_x\, x(\boldsymbol{y}, p_u; h) . \tag{4.25}$$

In addition, using that the amplitude $a'(\boldsymbol{y}, k_x, \omega; h)$ varies slowly over the spectral support (i.e., a spectral wedge) of a curvelet, we approximate

$$a'(\boldsymbol{y}, k_x, \omega; h) \approx a'(\boldsymbol{y}, k_x^u, \omega_u; h) . \tag{4.26}$$

Then, using equations (4.25) and (4.26) in (4.16), it follows that

$$\beta_\mu(\boldsymbol{y}) = a'(\boldsymbol{y}, k_x^u, \omega_u; h) \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \hat{c}_\mu(k_x, \omega) e^{-i\{\omega\, t(\boldsymbol{y}, x(\boldsymbol{y}, p_u; h); h) - k_x\, x(\boldsymbol{y}, p_u; h)\}} \mathrm{d}k_x \mathrm{d}\omega , \tag{4.27}$$

which is indeed simply an inverse Fourier transform of a curvelet.

Furthermore, we make use of the fact that curvelets remain (fairly) localized in the spatial domain under the action of the migration operator (see Figure 4.4 for an illustration in the context of time migration). That is, given a curvelet in the data $c_\mu(x, t)$ with center location (x_u, t_u) and main slope p_u, this curvelet is localized near the map-migrated location $\boldsymbol{y}_m$. Therefore, we linearize the dependence of $t(\boldsymbol{y}, x(\boldsymbol{y}, p_u; h); h)$ and $x(\boldsymbol{y}, p_u; h)$ on $\boldsymbol{y}$ around the map-migrated location $\boldsymbol{y}_m$. Linearizing $x(\boldsymbol{y}, p_u; h)$ gives

$$x(\boldsymbol{y}, p_u; h) \approx x_u + (\nabla_{\boldsymbol{y}} x(\boldsymbol{y}, p_u; h))|_{\boldsymbol{y}_m} \cdot (\boldsymbol{y} - \boldsymbol{y}_m) , \tag{4.28}$$

where we used $x_u = x(\boldsymbol{y}_m, p_u; h)$ and where the notation $\nabla_{\boldsymbol{y}}$ denotes the gradient with respect to $\boldsymbol{y}$. Linearizing $t(\boldsymbol{y}, x(\boldsymbol{y}, p_u; h); h)$ in $\boldsymbol{y}$ again involves total derivatives instead of

partial derivatives since both t and x depend on $\boldsymbol{y}$. Hence, we have

$$\begin{aligned} t(\boldsymbol{y},x(\boldsymbol{y},p_u;h);h) &\approx t(\boldsymbol{y}_m,x(\boldsymbol{y}_m,p_u;h);h) + \left.\frac{dt(\boldsymbol{y},x(\boldsymbol{y},p_u;h);h)}{d\boldsymbol{y}}\right|_{\boldsymbol{y}} \cdot (\boldsymbol{y}-\boldsymbol{y}_m) \\ &= t_u + (\boldsymbol{y}-\boldsymbol{y}_m)\cdot\left\{\nabla_{\boldsymbol{y}} t(\boldsymbol{y},x(\boldsymbol{y},p_u;h);h) + \frac{\partial t(\boldsymbol{y},x(\boldsymbol{y},p_u;h);h)}{\partial x}\nabla_{\boldsymbol{y}} x(\boldsymbol{y},p_u;h);h)\right\}\Big|_{\boldsymbol{y}_m} \\ &= t_u + (\boldsymbol{y}-\boldsymbol{y}_m)\cdot\left\{\nabla_{\boldsymbol{y}} t(\boldsymbol{y},x(\boldsymbol{y},p_u;h);h)|_{\boldsymbol{y}_m} + p_u\nabla_{\boldsymbol{y}} x(\boldsymbol{y},p_u;h);h)|_{\boldsymbol{y}_m}\right\} . \end{aligned} \tag{4.29}$$

Using equations (4.28) and (4.29) in (4.27) it follows that

$$\begin{aligned} \beta_\mu(\boldsymbol{y}) &= a'(\boldsymbol{y},k_x^u,\omega_u;h)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\hat{c}_\mu(k_x,\omega)e^{-i(\omega t_u - k_x x_u)} \\ &\quad \times e^{-i\left\{\omega\left\{\nabla_{\boldsymbol{y}} t(\boldsymbol{y},x(\boldsymbol{y},p_u;h);h)+p_u\nabla_{\boldsymbol{y}} x(\boldsymbol{y},p_u;h);h)\right\}|_{\boldsymbol{y}_m} - k_x\nabla_{\boldsymbol{y}} x(\boldsymbol{y},p_u;h)|_{\boldsymbol{y}_m}\right\}\cdot(\boldsymbol{y}-\boldsymbol{y}_m)}\mathrm{d}k_x\mathrm{d}\omega . \end{aligned} \tag{4.30}$$

Realizing that the integral is peaked at $\boldsymbol{y}=\boldsymbol{y}_m$, we can replace $a'(\boldsymbol{y},k_x^u,\omega_u;h)$ in equation (4.30) with $a'(\boldsymbol{y}_m,k_x^u,\omega_u;h)$. Doing this, while recognizing the inverse Fourier transform, we finally have

$$\beta_\mu(\boldsymbol{y}) = a'(\boldsymbol{y}_m,k_x^u,\omega_u;h)\; c_\mu(L\cdot(\boldsymbol{y}-\boldsymbol{y}_m)+\boldsymbol{x}_u) , \tag{4.31}$$

where we have defined

$$L := \left(\nabla_{\boldsymbol{y}} x(\boldsymbol{y},p_u;h)|_{\boldsymbol{y}_m} \quad \{\nabla_{\boldsymbol{y}} t(\boldsymbol{y},x(\boldsymbol{y},p_u;h);h)+p_u\nabla_{\boldsymbol{y}} x(\boldsymbol{y},p_u;h)\}|_{\boldsymbol{y}_m}\right)^T , \tag{4.32}$$

$$\boldsymbol{x}_u := \begin{pmatrix} x_u & t_u \end{pmatrix}^T . \tag{4.33}$$

Therefore, it follows that *the leading-order approximation (in angular frequency, horizontal wavenumber, and migrated location) to CO time-migration, consists of a coordinate transformation applied to a curvelet in the data*, given by

$$\boldsymbol{y} = L^{-1}\cdot(\boldsymbol{x}-\boldsymbol{x}_u)+\boldsymbol{y}_m , \tag{4.34}$$

combined with an amplitude scaling, where both the matrix L and the amplitude a' are evaluated with the use of pre-stack map migration. In other words, *Kirchhoff diffraction stacking is replaced with a coordinate transformation of the curvelets in the data.*

The matrix L can be written as

$$L = S_{p_u}\cdot T , \tag{4.35}$$

with

$$S_{p_u} := \begin{pmatrix} 1 & 0 \\ p_u & 1 \end{pmatrix}, \tag{4.36}$$

$$T := \left(\nabla_{\boldsymbol{y}} x(\boldsymbol{y}, p_u; h)|_{\boldsymbol{y}_m} \quad \nabla_{\boldsymbol{y}} t(\boldsymbol{y}, x(\boldsymbol{y}, p_u; h); h)|_{\boldsymbol{y}_m} \right)^T . \tag{4.37}$$

The explicit form of the matrix T for CO time-migration in (4.37) is derived in Appendix K. Note that the matrix S_{p_u} defines a unilateral shear along the t-axis; i.e., the matrix S_{p_u} shears the input curvelet with slope p_u along the t-axis to have zero slope.

Now that we know the leading-order approximation to one CO time-migrated curvelet c_μ through equation (4.31), we can determine the total image of the (scattered) data u_s using the reconstruction formula (4.1). Letting M denote the 2.5-D CO time-migration operator, we have

$$\beta(\boldsymbol{y}) = [M u_s](\boldsymbol{y}) = \sum_{\mu \in \mathcal{M}} (u_s, c_\mu) [M c_\mu](\boldsymbol{y}) = \sum_{\mu \in \mathcal{M}} (u_s, c_\mu) \beta_\mu(\boldsymbol{y}) , \tag{4.38}$$

where $\mathcal{M}$ is the index set containing all multi-indices μ represented in the data. Finally, because seismic data can be sparsely represented with curvelets, we can write the total image $\beta(\boldsymbol{y})$ as

$$\beta(\boldsymbol{y}) = \sum_{\mu \in \tilde{\mathcal{M}}} (u_s, c_\mu) \beta_\mu(\boldsymbol{y}) , \tag{4.39}$$

with $\beta_\mu(\boldsymbol{y})$ given by equation (4.31), and $\tilde{\mathcal{M}}$ the index set that holds the multi-indices of curvelets that survive a certain threshold. Note that the index set $\tilde{\mathcal{M}}$ is determined through thresholding the projection of the seismic data u_s onto the curvelet frame.

4.7.1 Input- versus output-based imaging

Equation (4.39) is based on a decomposition of the data $u_s(\boldsymbol{x})$ with curvelets, followed by thresholding of the resulting coefficients, and subsequent imaging of each curvelet that survived the thresholding. Therefore, this is an *input-based* seismic migration. This is analogous to convolution of each sample in the data with its associated isochron in accordance with classical Kirchhoff time-migration. Alternatively, *output-based* Kirchhoff migration can be calculated based on diffraction stacks, as is done in practical imaging algorithms; i.e., for each position in the image, a diffraction surface is calculated and the data are stacked over this surface. The resulting stack then gives the image at that particular position in the image.

There is an important difference between input- and output-based migration algorithms. In output-based algorithms the image points are specified and the needed diffraction surfaces in depth migration are calculated by ray-tracing in heterogeneous (anisotropic) media from the image point to the surface. In input-based algorithms, for a particular point in the data, all locations in the image that have the same traveltime as the data point need to be calculated. This is a more complex procedure than ray-tracing from the image point

to the subsurface. For this reason many practical migration implementations are output-based algorithms. For the homogeneous media case treated in this work, this difference is irrelevant. Formulating the migration as an output-based imaging algorithm, we have

$$\begin{aligned}\beta(\boldsymbol{y}) &= [Mu_s](\boldsymbol{y}) = \sum_{\mu' \in \mathcal{M}'} \left([Mu_s], c_{\mu'}\right) c_{\mu'}(\boldsymbol{y}) \\ &= \sum_{\mu' \in \mathcal{M}'} \left(u_s, [M^* c_{\mu'}]\right) c_{\mu'}(\boldsymbol{y}) \,, \end{aligned} \tag{4.40}$$

where M^* is the adjoint of the migration operator, i.e., the modeling (or demigration) operator. In this case, a curvelet in the image domain is demigrated and its projection onto the data is calculated. Subsequent thresholding then determines if the projection of the demigrated curvelet onto the data exceeds a specified threshold. This procedure would allow *target-oriented* wave-equation-based migration by demigrating curvelets contributing to a particular part of the image only. This requires calculating the leading-order approximation to the demigration operator acting on a single curvelet, i.e. $[M^* c_{\mu'}](\boldsymbol{x})$ in equation (4.40).

4.7.2 Analysis of the linear transformation L

In Appendix K we derive an explicit expression, equation (K.13), for the linear transformation T in 2D CO time-migration. It is given by

$$T = \begin{pmatrix} 1 & \dfrac{\tan\theta_s r_g^3 + \tan\theta_g r_s^3}{r_s^3 + r_g^3} \\ \dfrac{-2\cos\theta\sin\phi}{v} & \dfrac{2\cos\theta\cos\phi}{v} \end{pmatrix} , \tag{4.41}$$

where θ is the half opening-angle (see Figure 4.7). Inspection of the second row of this matrix, reveals that this transformation contains a rotation with angle $\phi = (\theta_s + \theta_g)/2$, which is the migrated dip. Making this rotation explicit, we rewrite T as

$$T = T' \cdot R_{-\phi} \,, \tag{4.42}$$

with

$$T' := \begin{pmatrix} \cos\phi + X\sin\phi & X\cos\phi - \sin\phi \\ 0 & \dfrac{2\cos\theta}{v} \end{pmatrix} , \quad R_{-\phi} := \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} . \tag{4.43}$$

The matrix $R_{-\phi}$ describes a rotation with angle ϕ. Because ϕ is clockwise positive and since t increases downward while x increases to the right, $R_{-\phi}$ describes an anti-clockwise rotation. In the definition of T' we introduced

$$X := \frac{\tan\theta_s r_g^3 + \tan\theta_g r_s^3}{r_s^3 + r_g^3} \tag{4.44}$$

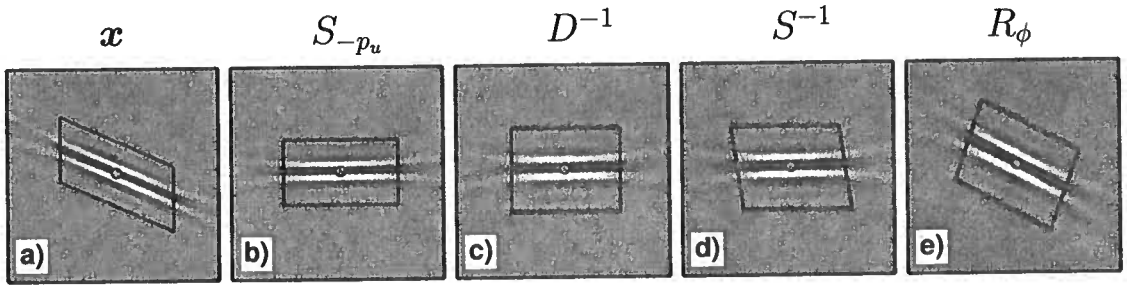


Figure 4.8. Illustration of the individual components of the linear transformation $L^{-1} = R_\phi \cdot S^{-1} \cdot D^{-1} \cdot S_{-p_u}$ [equation (4.47)]: input curvelet (a) after S_{-p_u} (b), D^{-1} (c), S^{-1} (d), and R_ϕ (e). The parallellogram superimposed on a curvelet in subfigure a and its transformed versions in b-e allow easy identification of the action of each individual component of L^{-1}.

for notational convenience. Further, inspecting equation (4.43) it follows that T' can be written as the matrix product of a dilation matrix D and a unilateral shear matrix S. That is, we can write T' as

$$T' = D \cdot S \,, \tag{4.45}$$

with

$$D := \begin{pmatrix} \cos\phi + X \sin\phi & 0 \\ 0 & \dfrac{2\cos\theta}{v} \end{pmatrix} \,, \quad S := \begin{pmatrix} 1 & \dfrac{X\cos\phi - \sin\phi}{\cos\phi + X\sin\phi} \\ 0 & 1 \end{pmatrix} \,. \tag{4.46}$$

Therefore, using equations (4.42) and (4.45) in equation (4.35), it follows that

$$L = S_{p_u} \cdot D \cdot S \cdot R_{-\phi} \,. \tag{4.47}$$

Finally, using this representation of L in the coordinate transformation in equation (4.34), which describes the leading-order contribution to 2D CO time-migration, we have

$$\boldsymbol{y} = R_\phi \cdot S^{-1} \cdot D^{-1} \cdot S_{-p_u} \cdot (\boldsymbol{x} - \boldsymbol{x}_u) + \boldsymbol{y}_m \,, \tag{4.48}$$

where we used that $S_{-p_u} = S_{p_u}^{-1}$.

From equation (4.48) it follows that the leading-order approximation to CO time-migration can be described by the following sequence of linear[10] transformations. First, a curvelet in the data with center location $\boldsymbol{x}_u$ (i.e., center midpoint location x_u and center two-way traveltime t_u) and center slope p_u, is sheared along the time-axis to have zero slope. Subsequently, the curvelet is dilated in both the vertical and horizontal direction. Then, the curvelet is sheared along the horizontal direction and is rotated to have output dip ϕ.

[10]The translation part of the coordinate transformation is of course not linear.

All these linear transformations have the center location $\boldsymbol{x}_u$ of the curvelet as their origin. Finally, the resulting transformed curvelet is translated to the migrated output location $\boldsymbol{y}_m$. This sequence of transformations (except the translation) is depicted for one curvelet in Figure 4.8, where a parallellogram is added to illustrate the geometric character of these transformations.

Based on our knowledge of map migration, the rotation and translation part of the sequence of transformations is intuitive. The dilation in the vertical direction with $v/(2\cos\theta)$ (after shearing the curvelet along the time axis to have zero slope) takes care of both the time-depth conversion and of the lowering of the frequency content resulting from the imaging condition; i.e., the frequency content is lowered by a factor $\cos\theta$ because the length of the migrated slowness $\boldsymbol{p}_m$ is smaller than the sum of the lengths of the input slownesses $\boldsymbol{p}_{s,r}$ (see Figure 4.7). Therefore, the translation, rotation and dilation part of the transformation conform to our intuition [see also Douma & de Hoop (2005)]. However, the unilateral shears along both the vertical and horizontal axes, as well as the dilation of the horizontal axis, are somewhat surprising, and indicate that the leading-order approximation to CO time-migration incorporates a certain amount of linear deformation — this linearity results from the linearization of $t(\boldsymbol{y}, x(\boldsymbol{y}, p_u; h); h)$ and $x(\boldsymbol{y}, p_u; h)$ about $\boldsymbol{y}_m$. To understand this, consider the special case of zero-offset migration. In that case the matrices S_{-p_u} and R_ϕ remain unchanged, while the shear matrix S becomes the identity matrix and the dilation matrix D is given by

$$D|_{h=0} = \begin{pmatrix} 1/\cos\phi & 0 \\ 0 & 2/v \end{pmatrix}. \tag{4.49}$$

Observe that this matrix includes no $\cos\theta$ term because for zero-offset $\theta = 0$ (i.e., $|\boldsymbol{p}_m| = 2/v$). Note also that this matrix includes a squeeze in the horizontal direction equal to $\cos\phi$. Because this squeeze is applied after S_{-p_u}, which shears the input curvelet to have zero slope, the squeeze shortens the long axis of a curvelet. If $\phi = \pi/2$ the long axis of the curvelet is mapped onto a point. This can be understood by noticing that, assuming a constant velocity, we can record a reflection from a dipping reflector with dip $\pi/2$ in the subsurface only if this reflector is at the surface. Since the reflection of this reflector will have a slope $2/v$ in the data, a straight line with this slope will be mapped onto a point. Therefore the horizontal component of the dilation matrix accounts for both focussing (or defocussing) of curvelets upon imaging and geometrical spreading.

4.7.3 Calculation of L

All quantities necessary for the calculation of L, can be found from the curvelet decomposition of the data, combined with map migration. First, the central location $\boldsymbol{x}_u = (x_u\ t_u)^T$ of the curvelet in the data is found from the translation indices m_1 and m_2, while the slope p_u is found from the scale index j and the angular index l. This slope determines the matrix S_{-p_u}. Then, using x_u, t_u, and p_u in the 2D map time-migration equations (4.2)–(4.4), we find the migrated location $\boldsymbol{y}_m$, and the migrated slope p_m. The migrated dip ϕ can then be calculated from $\phi = \tan^{-1}(vp_m/2)$, which determines the rotation matrix R_ϕ. From $\boldsymbol{y}_m$ and x_u the distances $r_{s,g}(\boldsymbol{y}_m, x_u; h)$ can be calculated using equation (4.9) for a given

half-offset h, and the angles $\theta_{s,g}$ then follow from $\theta_{s,g} = \cos^{-1}(y_2^m/r_{s,g})$, which determine the half opening-angle $\theta = (\theta_g - \theta_s)/2$. Using the calculated values of ϕ, $r_{s,g}(\boldsymbol{y}_m, x_u; h)$, $\theta_{s,g}$, and θ, we can calculate the matrices D and S using equations (4.44) and (4.46). Note that in digital implementations the matrix L needs to be applied on a grid. For this purpose, we derive the discrete form of L in Appendix L.

4.8 Extension to pre-stack depth-migration

Although here we have concerned ourselves with time-migration only (i.e., straight rays), the main ingredients of the presented derivation apply to pre-stack depth-migration as well. That is, for pre-stack depth-migration the derivation of the leading-order approximation also contains a stationary phase evaluation of the integral with respect to the non-constant (or non-common) variable[11], and subsequent linearization of the phase of the resulting oscillatory integral followed by linearization around the (map) migrated location $\boldsymbol{y}_m$. Therefore, for pre-stack depth-migration, a derivation similar to that presented here also leads to the leading-order approximation to depth-migration being a simple transformation of coordinates. For depth migration, however, the transformation is calculated for a smoothed medium, as pointed out by Smith (1998) in the context of solving the wave-equation; i.e., a curvelet with scale index j (i.e., a wavelength proportional to 2^{-j}) is migrated through a smoothed medium that has no heterogeneity beyond a scale proportional to $\sqrt{2^{-j}} = 2^{-j/2}$, i.e., beyond the width of a curvelet. This is reminiscent of the first Fresnel zone of a wave being proportional to $\sqrt{\lambda}$, with λ the wavelength.

In the context of one-way wave-equation-based migration we mention that deformation of a curvelet beyond the leading-order approximation could possibly be calculated using an approach similar to the one given by Smith (1998, p.799), who presents a weak solution to the wave equation as the solution of a Volterra equation. The solution of this equation is evaluated based on an initial estimate of the solution constructed from roughly translating the centers of curvelets along the Hamiltonian flow associated with the smoothed media. In the context of one-way wave-equation-based imaging, the wave operator is to be replaced by a one-way operator, while time is replaced by depth. Then the extension of the presented leading-order approximation for CO time-migration to pre-stack depth-migration, would allow the calculation of the initial estimate of the solution of the Volterra equation. We emphasize that this idea remains to be verified.

4.9 Numerical examples

Figure 4.9a shows the 2.5-D CO Kirchhoff migrated curvelet from Figure 4.4 in more detail, while Figure 4.9d shows the real part of the associated spectrum. Figure 4.9b shows the result of using the leading-order approximation (4.31) to image the same curvelet, while Figure 4.9c shows the difference between the Kirchhoff result and that of the leading-order

[11] In this work the non-common variable is the midpoint, but in, for example, common-shot depth-migration, it would be the receiver location.

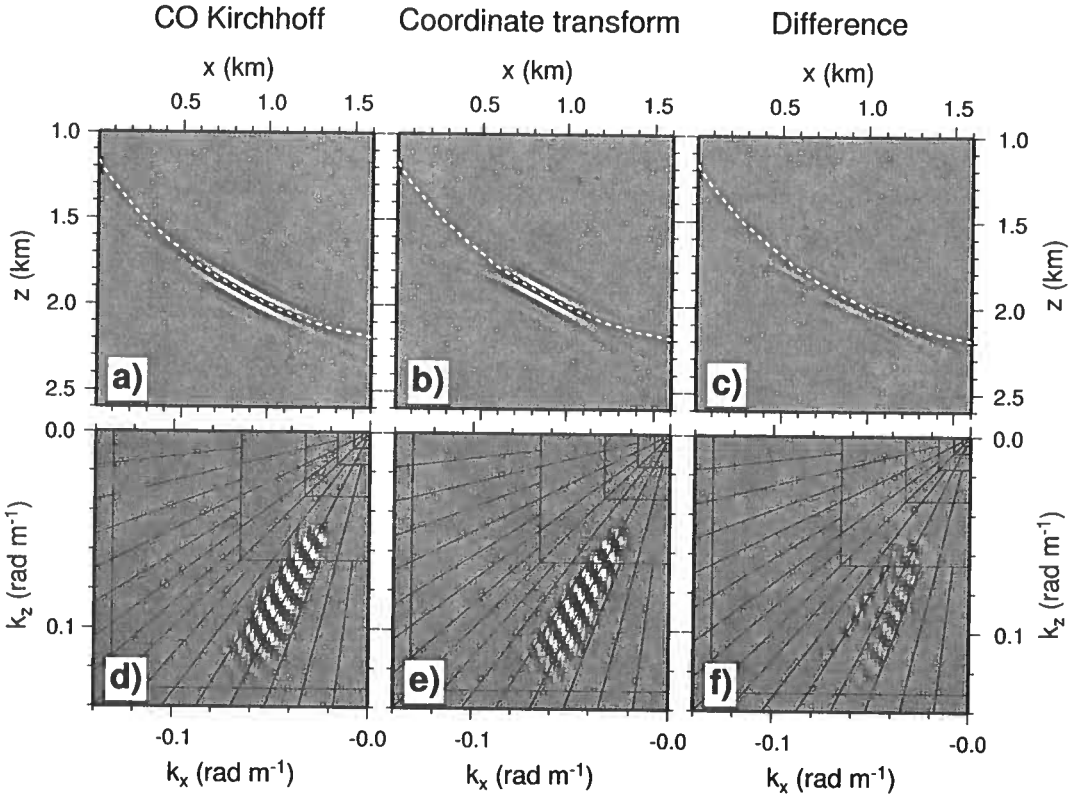


Figure 4.9. CO Kirchhoff migration of the curvelet shown in Figure 4.4 (a), coordinate transformation of this curvelet according to equation (4.34) (b), and the difference (c). The bottom row shows the real part of the spectrum for the CO Kirchhoff migrated curvelet (d), the coordinate transform of equation (4.34) (e), and the difference (f). The coordinate transformation provides a good first approximation to the Kirchhoff-migrated curvelet.

approximation. Before subtraction, both images were normalized to the same maximum amplitude so that the difference shows only relative amplitude differences between the two images. Figures 4.9e and f show the real part of the amplitude spectra of the images shown in Figures 4.9b and c.

The leading-order approximation (4.31) based on a simple transformation of coordinates of curvelets provides a good approximation to the Kirchhoff result. The maximum amplitude of the difference between both methods is 24% of the maximum amplitude in the Kirchhoff image. Note from the patterns in the real parts of the spectra, that the curvelet is slightly curved due to the migration, whereas the linear coordinate transformation does not take such bending into account (cf., Figures 4.9d and e). As a consequence, the main difference in the spectrum occurs on the edges of the support of the curvelet in the frequency domain (see Figure 4.9f). This difference, attributable to the nonlinear deformation of the curvelet which is absent in the presented leading-order approximation, can again be considered small.

Although Figure 4.9 shows a favorable comparison for one particular curvelet, of importance is the interference among several curvelets after migration. Figure 4.10a shows a superposition of eight curvelets, with the same central location in space, and the same scale index, but different directions (or angular indices). Figure 4.10b shows the amplitude spectrum of the superposition of all eight curvelets, revealing that we used right-ward sloping curvelets only. Figure 4.10c shows the 2.5-D CO Kirchhoff migrated result (the offset is here 2 km), while Figure 4.10d shows the result of the leading-order approximation [i.e., equation (4.38), with $\beta_\mu(\boldsymbol{y})$ calculated using (4.31)]. Since the data in Figure 4.10a contain right-ward sloping curvelets only, the Kirchhoff result contains the left part of the bandlimited isochron only, as expected. Comparing Figures 4.10c and d, the leading-order approximation again provides a good approximation to the Kirchhoff migrated result; the interference of the different curvelets leads to an overal smooth left-part of the bandlimited isochron that compares favorably with the Kirchhoff result. Away from the isochron, both results show the weak tails of the curvelets. For the Kirchhoff result these tails are again somewhat curved towards the isochron, while the leading-order approximation does not include such nonlinear deformation. The difference between both results, however, is small.

Figure 4.11a shows a CO section (with offset equal 1 km) of synthetic data generated from a syncline model in a constant-velocity background. We use these data to further test the curvelet-based migration, and to highlight that triplications in the data are not a problem; crossing events in the data are simply constructed from curvelets that have the same location but different orientations. In Figure 4.11a we used only the 0.75% largest curvelet coefficients of the data to reconstruct the data. This thresholding percentage was determined in the same manner as presented earlier for the data shown in Figure 4.5, i.e., through visually verifying that the difference between the original and reconstructed data is negligible. The associated amplitude spectrum of the reconstructed data is shown in Figure 4.11b. Observe that almost all energy is contained in frequency bands associated with scale indices $4 \leq j \leq 6$. Figure 4.11c shows the 2.5-D CO Kirchhoff migrated result, while Figure 4.11d shows the curvelet-based migration calculated with the leading-order approximation [i.e, using equation (4.39)]. The leading-order curvelet-based migration provides a good

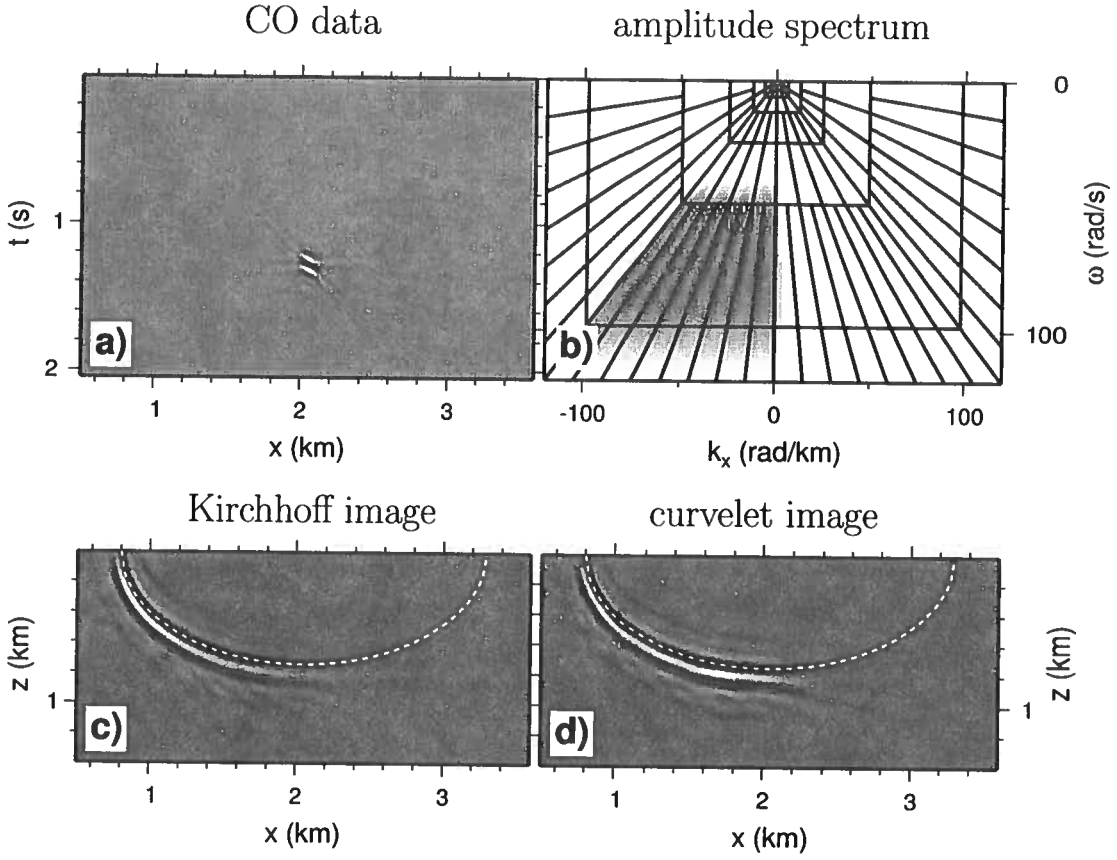


Figure 4.10. Superposition of eight curvelets (a), the associated amplitude spectrum (b), the CO Kirchhoff-migrated result (c), and the result from the coordinate transformation of equation (4.34), combined with amplitude scaling [equation (4.31] (d). The amplitude-corrected and coordinate-transformed curvelet gives a good approximation to the Kirchhoff-migrated result, and the interference between different curvelets results in a smooth isochron. Since we used only rightward-sloping curvelets, the left part of the isochron is constructed only.

ures 4.12a, b and c). Figures 4.12d, e, and f, show the associated cumulative images; the images for $j < 4$ are not shown because the data have essentially no energy for $j < 4$ (see Figure 4.11). Following such per-scale imaging, the associated image gathers can then be submitted to a velocity analysis. The opportunity for such frequency-dependent migration velocity analysis is a straightforward consequence of the curvelet frame being an appropriate representation of both the seismic data and the imaging operator. Note that per-scale imaging also allows the study of the finite-frequency interaction with the interfaces, i.e., the frequency-dependence of the reflection coefficient.

The locally-associated main directions of curvelets allow controlled illumination (Rietveld & Berkhout, 1992) of the subsurface. Figure 4.13a shows the amplitude spectrum of the data in Figure 4.11a, but reconstructed with rightward sloping curvelets only. Hence, a dip-filtering is here achieved by simply reconstructing the data using rightward sloping curvelets only. Figure 4.11b shows the resulting image, which now mainly shows the left part of the syncline only. Note that a partial image of the discontinuity on the right edge of the syncline is also visible, but now imaged with rightward sloping curvelets only. Here, the controlled illumination is achieved by dip-filtering the *data* with curvelets. Dip-filtering the *image* with curvelets, in contrast, leads naturally to the focusing-in-dip procedure of Brandsberg-Dahl *et al.* (2003). The output-based imaging algorithm based on equation (4.40) would then simply demigrate only curvelets in the image with certain dips in certain locations, and check if the resulting demigrated curvelet had a projection on the data that exceeded a chosen threshold. Again, these opportunities are in essence a straightforward consequence of curvelets being appropriate building blocks of both seismic data and the imaging operator; no additional preprocessing, such as slant-stacking, is needed.

4.10 Discussion

Here, we have calculated the coordinate transformation in equation (4.34) using a brute-force approach in the spatial domain. For each significant curvelet coefficient, we apply an inverse curvelet transform, and transform the resulting curvelet in the spatial domain according to equation (4.34). This approach has allowed us to generate the numerical examples and to show the proof of concept of imaging with curvelets using the derived leading-order approximation. This approach, however, does not provide an efficient algorithm to perform seismic imaging with curvelets. Ultimately one would want to calculate the transformation of coordinates in the curvelet frame through a mapping of curvelet indices and coefficients, although an approach that makes use of the compact support of the curvelets in the spectral domain, seems worth investigating also.

Curvelet-based seismic imaging can have several potential benefits over existing imaging algorithms. As mentioned in the previous section, the fact that curvelets incorporate the character of the seismic data combined with the fact that they are elements of phase-space, turns extensions to seismic imaging such as controlled illumination, focussing in dip, and frequency-dependent migration velocity analysis, into natural consequences of an appropriate reparameterization of seismic data and the imaging operator. Moreover, the fact that the leading-order approximation to curvelet-based seismic imaging is calculated using map

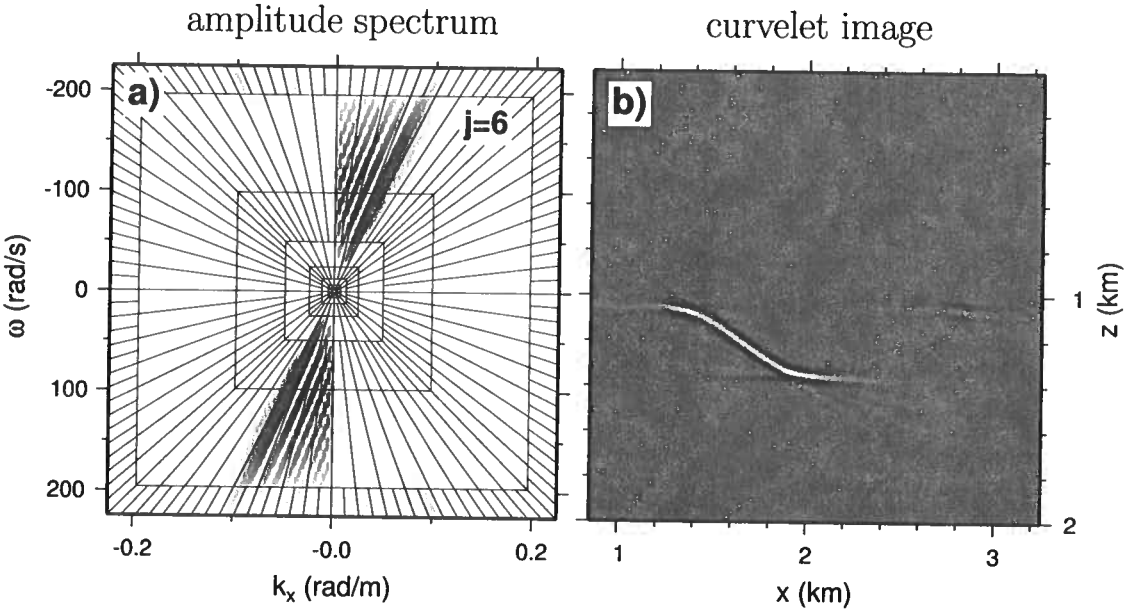


Figure 4.13. Amplitude spectrum of the thresholded data in Figure 4.11a, but now using only rightward-sloping curvelets (a), and the resulting image of these curvelets using equation (4.39) (b). By using curvelets with certain slopes (i.e., certain angular indices) only, we achieve a controlled illumination of the subsurface.

migration, suppresses the need for integration over diffraction surfaces. Combined with the intuition that curvelets allow a sparse representation of the data, which we exemplified here using simple examples, this could therefore lead to a gain in computation time over existing imaging algorithms. However, we currently do not know how much additional computational overhead is needed to calculate the deformation of curvelets beyond the leading-order approximation. Therefore, the question whether curvelet-based imaging will lead to a gain in computational efficiency over existing algorithms remains currently open. Finally, based on the work of Smith (1998), curvelet-based imaging allows imaging through media with limited smoothness, i.e., rougher media.

Even though the curvelets used in this paper are two-dimensional, they can be extended to higher dimensions (Ying *et al.*, 2005). Figure 4.14 shows an example of a three-dimensional (3D) curvelet in both the spatial domain (a) and the spectral domain (b). In the spatial domain, 3D equivalents of curvelets look like circular disks that are smooth along the disk and oscillatory orthogonal to the disk. Roughly speaking they are smoothed circular pieces of a bandlimited plane wave in 3D[12]. Hence, provided the leading-order approximation to seismic imaging is worked out in three dimensions in, e.g., a common-source setting (or common-offset as in this work), 3D curvelets allow seismic imaging in 3D.

4.11 Conclusion

With curvelets as building blocks of seismic data, the character of the data, i.e., bandlimitation and locally associated directions, can be built into the representation of the data. A simple projection of the data onto the curvelet frame, combined with an intelligent thresholding algorithm, then provides the local slopes in the data. We have presented simple numerical examples to illustrate this. These examples also illustrate the potential of curvelets to sparsely represent the data, even though this needs to be verified for field data in combination with a more sophisticated denoising algorithm.

We have shown that using curvelets, the leading-order approximation to CO time-migration becomes a simple transformation of the input coordinates of the curvelets in the data, combined with amplitude scaling. In other words, to leading order in angular frequency, horizontal wavenumber, and migrated location, the Kirchhoff diffraction stack can be replaced with a coordinate transformation of curvelets in the data. This transformation is calculated using CO map time-migration, which uses the local slopes provided by the curvelet decomposition of the data. Considering that the data can be sparsely represented with relatively few curvelets, CO time-migration becomes to leading order a transformation of coordinates of only few curvelets combined with amplitude scaling. We have presented numerical examples indicating that this leading-order approximation provides a good approximation to such time migration.

Although the current work treats constant media only, we emphasize that the essence of the presented derivation applies to pre-stack depth-migration also. Specifically, the

[12]This rough description ignores that each curvelet has a small range of $\boldsymbol{k}$-vectors associated with it, rather than only a single $\boldsymbol{k}$ direction, as has a plane wave.

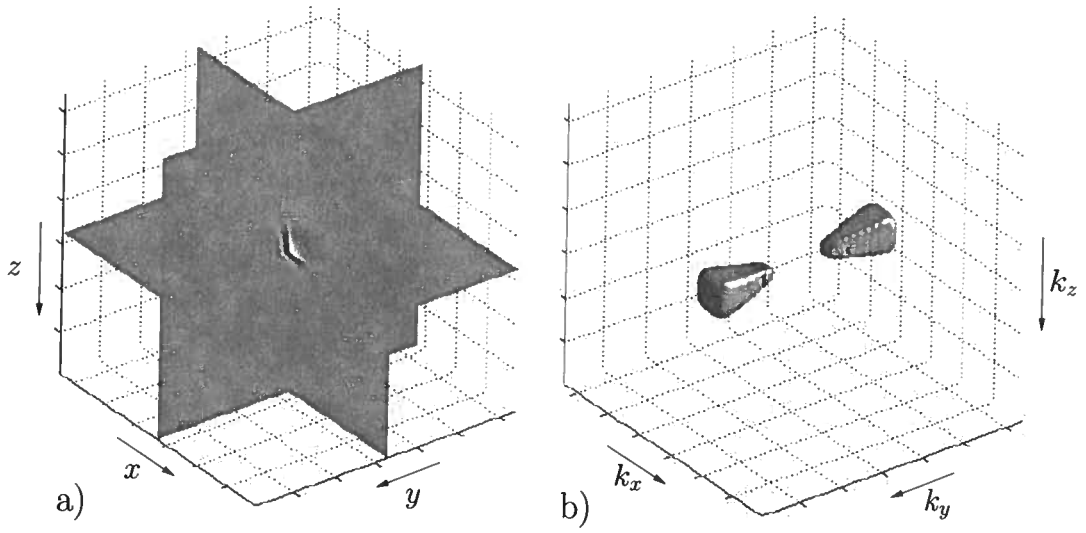


Figure 4.14. A 3D curvelet in space (a) and its associated (contoured) amplitude spectrum (b). In the spatial domain, 3D equivalents of curvelets look like circular disks that are smooth along the disk and oscillatory orthogonal to the disk.

derivation of the leading-order approximation to pre-stack depth-migration also contains a stationary phase evaluation of the integral with respect to the non-constant variable, and subsequent linearization of the phase of the resulting oscillatory integral followed by linearization around the map-migrated location of the center of the curvelet. In addition, we briefly sketched a possible way to account for deformation of the curvelet beyond the leading-order approximation while emphasizing that this idea remains to be verified in the future. Because curvelets are basically elements of phase-space, curvelet-based seismic imaging naturally allows the formation of caustics in heterogeneous anisotropic media. Given the bandlimited nature of curvelets, curvelet-based migration combines high-frequency-asymptotics-based Kirchhoff imaging with finite-frequency-based wave-equation migration. This work is in our opinion the first of a series of papers that will possibly lead to a new generation of seismic imaging algorithms.

Chapter 5

Nonhyperbolic moveout analysis of qP-waves in VTI media using rational interpolation[1]

5.1 Summary

Anisotropic velocity analysis using qP-waves in transversely isotropic media with a vertical symmetry axis (VTI), is usually done by inferring the anellipticity parameter η and the normal moveout velocity V_{NMO} from the nonhyperbolic character of the moveout. Several approximations explicit in these parameters exist with varying degrees of accuracy. Here, we present a rational interpolation approach to nonhyperbolic moveout analysis in the $t-x$ domain. This method has no additional computational overhead compared to using expressions explicit in η and V_{NMO}. The lack of such overhead stems from the observation that, for fixed η and zero-offset two-way traveltime t_0, the moveout curve for different values of V_{NMO} can be calculated by simple stretching of the offset axis. This observation is based on the assumptions that the traveltimes of qP-waves in (laterally homogeneous) VTI media, depend mainly on η and V_{NMO}, and that the shear-wave velocity along the symmetry axis has a negligible influence on these traveltimes. The accuracy of the rational interpolation method is as good as that of these approximations. The method can be tuned to be accurate to any offset range of interest, by increasing the order of the interpolation.

We test the method using both synthetic and field data, and compare it with the nonhyperbolic moveout equation of Alkhalifah & Tsvankin (1995), and the shifted hyperbola equation of Fomel (2004). Both data types confirm that for $\eta \gtrsim 0.1$ our method significantly outperforms the nonhyperbolic moveout equation in terms of combined unbiased parameter estimation with accurate moveout correction. Comparison with the shifted hyperbola equation of Fomel for Greenhorn-shale anisotropy establishes almost identical accuracy of the rational interpolation method and his equation. Even though the proposed method currently deals with homogeneous media only, results from application to synthetic and field data confirm the applicability of the proposed method to horizontally layered VTI media.

5.2 Introduction

Over the past two decades, the importance of anisotropy and its influence on seismic data processing have become increasingly appreciated. Since 75% of the clastic infill of

[1]This Chapter has been accepted for publication in *Geophysics* as: Huub Douma and Alexander Calvert, 2006, Nonhyperbolic moveout analysis in VTI media using rational interpolation.

sedimentary basins consists of shale formations (Tsvankin, 2001, p.11), and since the transversely isotropic (TI) model adequately describes the intrinsic anisotropy of shales (Sayers, 1994), wave propagation in TI media has attracted much attention. Because the dispersion relations govern the propagation velocities of the different wave modes (and hence the traveltimes used in seismic data processing), and because these relations are nonlinear in the elastic coefficients, many authors have worked on approximations of the dispersion relations in TI media (Dellinger *et al.*, 1993; Tsvankin & Thomsen, 1994; Alkhalifah, 1998; Schoenberg & de Hoop, 2000; Zhang & Uren, 2001; Stovas & Ursin, 2004; Fomel, 2004), with varying levels of accuracy. Fowler (2003) gives a comparative review of some of these approximations with an emphasis on TI media with a vertical symmetry axis (VTI media).

The nonhyperbolic moveout equation for qP-waves developed by Tsvankin & Thomsen (1994) was derived to allow velocity analysis in VTI media using such waves. This equation was later rewritten by Alkhalifah & Tsvankin (1995) in terms of the normal moveout velocity V_{NMO} and the anellipticity parameter η, i.e., in terms of the relevant parameters for the kinematics of qP-waves in VTI media. Even though this equation is exact at zero offset and infinite offset, Grechka & Tsvankin (1998) mention that at intermediate offsets "this approximation can be somewhat improved by empirically changing the denominator of the nonhyperbolic term." This limited accuracy at intermediate offsets was also noted by other authors [e.g., Zhang & Uren (2001), Van der Baan & Kendall (2002), and Stovas & Ursin (2004)], and causes (as we show in this paper) a *bias* in the estimated value of η. As larger offset-to-depth ratios (ODR) provide better resolution for η (Alkhalifah, 1997; Grechka & Tsvankin, 1998; Wookey *et al.*, 2002), improved accuracy for larger ODR can help reduce the uncertainty in inversion for η [within the limits of the trade-off between η and V_{NMO}, as observed by Grechka & Tsvankin (1998)] when large ODR are available.

To achieve higher accuracy at larger ODR and to reduce the bias in the estimated values of η when using the nonhyperbolic moveout equation from Alkhalifah and Tsvankin, we propose a rational interpolation scheme for traveltimes of qP-waves in homogeneous VTI media. The choice to use a rational interpolation was motivated by the observation that several of the aforementioned approximations achieve high accuracy through the use of rational approximations [e.g., Schoenberg & de Hoop (2000) and Stovas & Ursin (2004)], and by the form of the nonhyperbolic moveout equation of Tsvankin & Thomsen (1994), which resembles a rational approximation. We refrain from an attempt to derive yet another approximation to traveltimes of qP-waves in such media that is explicit in the relevant parameters. Instead, we simply make use of the facts that the shear-wave velocity along the (vertical) symmetry axis, V_{S0} has negligible influence on the traveltimes of qP-waves in TI media (Tsvankin & Thomsen, 1994; Tsvankin, 1996; Alkhalifah, 1998, 2000), and that these traveltimes depend mainly on the anellipticity parameter η and the (zero-dip) normal-moveout velocity V_{NMO} (Alkhalifah & Tsvankin, 1995). That is, the influence of Thomsen parameter δ is only small. We show that these two assumptions cause the influence of V_{NMO} on the nonhyperbolic moveout of qP-waves in a horizontal homogeneous VTI layer, for fixed anellipticity parameter η and two-way zero-offset traveltime t_0, to be a simple horizontal stretch (or squeeze); i.e., the moveout curve is stretched (or squeezed) along the offset axis. This observation allows the traveltimes needed for the rational interpolation to be

calculated from a small number of precomputed traveltimes stored in a table. Therefore, this method has no additional computational overhead compared to methods that use explicit approximations in the relevant parameters.

The outline of this paper is as follows. First, we study the accuracy of the nonhyperbolic moveout equation from Tsvankin & Thomsen (1994). Subsequently, we introduce the rational interpolation approach to nonhyperbolic moveout for qP waves in VTI media, and explain the stretch-squeeze influence of V_{NMO} on the moveout curve of qP reflections in a single horizontal homogeneous VTI layer. Synthetic tests, for both a single homogeneous horizontal VTI layer and a horizontally layered VTI medium, subsequently verify the improved accuracy of the method when compared to the nonhyperbolic moveout equation. Finally these findings are confirmed by an application of the method to field data.

5.3 Accuracy of the nonhyperbolic moveout equation

The nonhyperbolic moveout equation of Tsvankin & Thomsen (1994), rewritten in terms of the anellipticity parameter η by Alkhalifah & Tsvankin (1995), is given by

$$t^2(x) = t_0^2 + \frac{x^2}{V_{NMO}^2} - \frac{2\eta x^4}{V_{NMO}^2\left[t_0^2 V_{NMO}^2 + (1+2\eta)\,x^2\right]}\,, \tag{5.1}$$

where t_0 is the two-way traveltime at zero offset, x is offset, and V_{NMO} is the (zero-dip) NMO velocity. Note how the moveout reduces to hyperbolic moveout for elliptical anisotropy ($\eta = 0$) (e.g., (Tsvankin, 2001, section 4.3.1). This equation is exact at both zero offset and infinite offset.

Figure 5.1a show contours of the maximum absolute traveltime difference (in percentage of t_0) between ray-traced traveltimes for a single horizontal VTI layer and the nonhyperbolic moveout equation, for virtually all combinations of η and V_{NMO} of practical interest. The maximum ODR is two, and $t_0 = 1$ s for all models considered. The ray-traced traveltimes were determined with $V_{S0} = 0$ km/s and $\delta = 0$. The differences in the traveltimes are of the order of a percent, which in this case amounts to 10 ms. For a dominant frequency of 50 Hz in surface seismic data, this amounts to a traveltime error of about half a dominant period; i.e., a sizeable error that can lead to a bias in the estimated values of η and V_{NMO} when velocity analysis is done using this equation. Note that here we use the true values of η and V_{NMO} for the moveout analysis rather than the best-fit values, since we want to analyze the accuracy of the nonhyperbolic moveout equation for a known model. It is known that the traveltime differences shown in Figure 5.1a can be reduced by using the best-fit values of η (and V_{NMO}), rather than the true ones; that is, the traveltime differences become smaller when a bias is introduced in the estimates of η (and V_{NMO}).

Figures 5.1b is as Figure 5.1a except that here the maximum ODR is four. With current acquisition systems, such values of ODR are feasible for shallow targets and are common in near-surface seismic applications. Note that for this ODR, the (relative) errors in the traveltimes have increased roughly with a factor two.

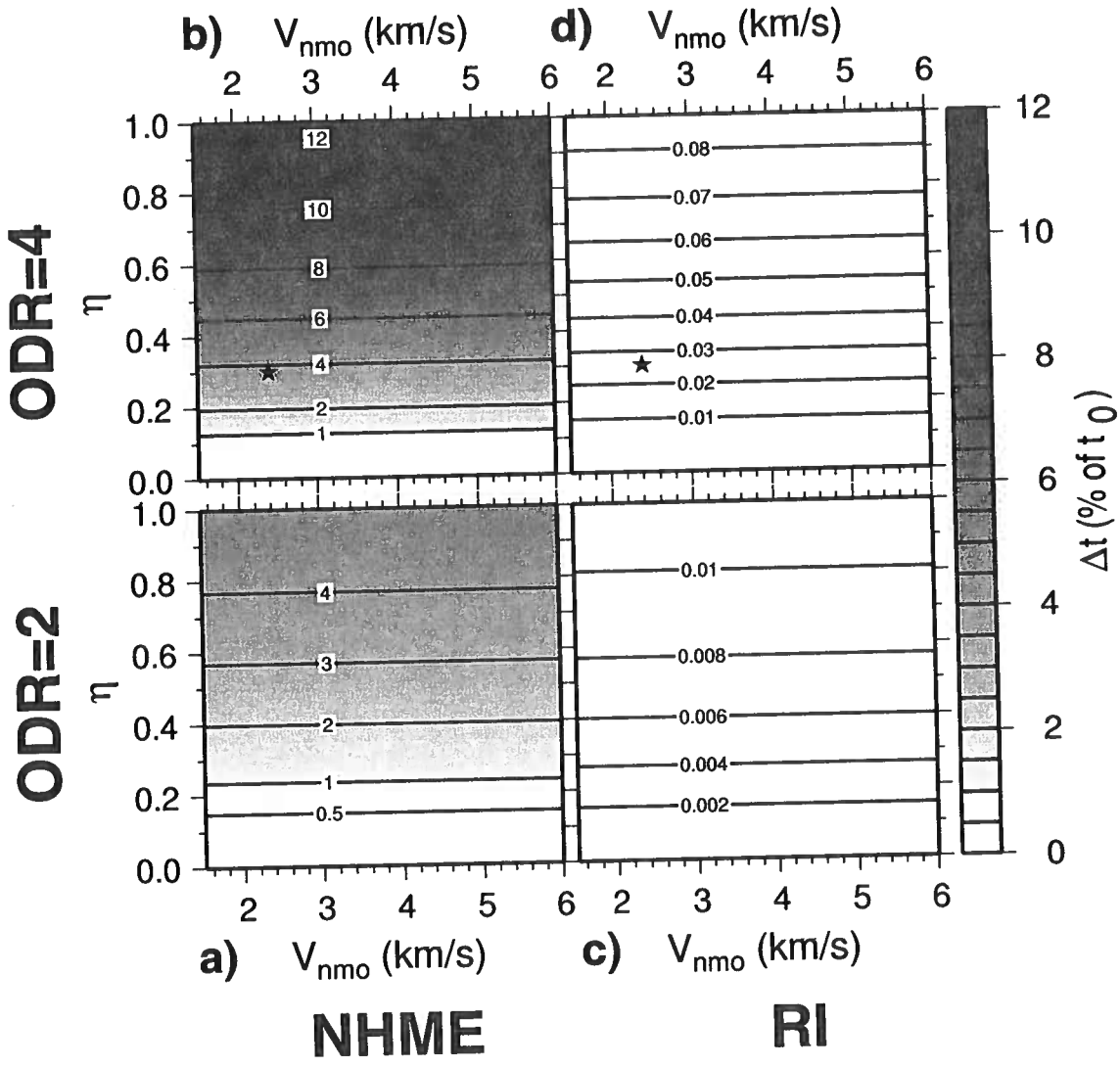


Figure 5.1. Comparison of the accuracy of the nonhyperbolic moveout equation (a and b, NHME) and of the [2/2] rational interpolation (c and d, RI), for a range of models (combinations of η and V_{NMO}) that spans most models of practical interest. In subfigures a and c the maximum ODR is two, whereas in b and d the maximum ODR is four. Contours are drawn for the maximum absolute percentage error in traveltime (compared to ray-traced traveltimes where ray tracing was done with $V_{S0} = 0$ km/s and $\delta = 0$) for a single horizontal VTI layer. The contoured values are shown in percentage of t_0. The stars in subfigures b and d indicate the model parameters used for the residual moveout plots in Figure 5.3.

5.4 The dependence of nonhyperbolic moveout for a single VTI layer on η and V_{NMO}

From simple geometric considerations, it follows that for a homogeneous horizontal VTI layer, the traveltime is given by

$$t = \frac{V_{P0}\, t_0}{v \cos\psi}\,, \tag{5.2}$$

where V_{P0} is the P-wave velocity along the vertical symmetry axis, ψ is the group angle, and v is the group velocity for propagation in direction ψ. The offset x associated with this traveltime is given by

$$x = V_{P0}\, t_0 \tan\psi\,. \tag{5.3}$$

In a TI medium, the phase velocity is given by (Tsvankin, 2001, p.22)

$$V(\theta) = V_{P0}\sqrt{1 + \epsilon \sin^2\theta - \frac{f}{2}\left(1 - \sqrt{\left(1 + \frac{2\epsilon\sin^2\theta}{f}\right)^2 - \frac{2\,(\epsilon-\delta)\sin^2 2\theta}{f}}\right)}\,, \tag{5.4}$$

where θ is the phase angle, ϵ and δ are the Thomsen parameters, and $f \equiv 1 - V_{S0}^2/V_{P0}^2$, with V_{S0} the S-wave phase velocity along the symmetry axis. For such a medium, the group angle is related to the phase angle through (Tsvankin, 2001, p.29)

$$\tan\psi = \frac{\tan\theta + \dfrac{1}{V(\theta)}\dfrac{dV}{d\theta}}{1 - \dfrac{\tan\theta}{V(\theta)}\dfrac{dV}{d\theta}}\,, \tag{5.5}$$

while the group velocity v is given by

$$v = V(\theta)\sqrt{1 + \left(\frac{1}{V(\theta)}\frac{dV}{d\theta}\right)^2}\,. \tag{5.6}$$

It is known that the qP-wave phase-velocity in TI media depends only weakly on V_{S0} (Tsvankin & Thomsen, 1994; Tsvankin, 1996; Alkhalifah, 1998). For all kinematic problems regarding qP waves, V_{S0} is therefore usually ignored. We likewise set $V_{S0} = 0$ (or $f = 1$) in equation (5.4). This is the *acoustic approximation* from Alkhalifah (1998, 2000). Note that Alkhalifah & Tsvankin (1995) obtained equation (5.1) from the nonhyperbolic moveout equation of Tsvankin & Thomsen (1994), by also setting $V_{S0} = 0$. In addition, Alkhalifah & Tsvankin (1995) showed that the time signatures of qP-waves in homogeneous VTI media depend mainly on the (zero-dip) normal-moveout velocity V_{NMO} and the anellipticity parameter η, with an almost negligible influence of V_{P0}. Since we are interested only in traveltime calculations, we can choose $\delta = 0$, and thus $V_{P0} = V_{NMO}$ and $\epsilon = \eta$, in equation

(5.4). That equation then becomes

$$V(\theta) = V_{NMO}\sqrt{\eta \sin^2\theta + \frac{1}{2}\left(1 + \sqrt{\left(1 + 2\eta\sin^2\theta\right)^2 - 2\eta\sin^2 2\theta}\right)}\,, \tag{5.7}$$

while equations (5.2) and (5.3) for the traveltime t and the associated offset x, become

$$t = \frac{V_{NMO}\, t_0}{v\cos\psi}\,, \tag{5.8}$$

and

$$x = V_{NMO}\, t_0 \tan\psi\,, \tag{5.9}$$

respectively. Note that the phase velocity $V(\theta)$ now depends linearly on V_{NMO}. This linearity causes the term $\frac{1}{V(\theta)}\frac{dV}{d\theta}$ in equations (5.5) and (5.6) to be independent of V_{NMO}. Since the dependence of the group angle on the anisotropic parameters is governed by the term $\frac{1}{V(\theta)}\frac{dV}{d\theta}$ [cf. equation (5.5)], the group angle ψ is independent of V_{NMO} and depends only on η. In addition, it follows from equation (5.6) that the group velocity v depends linearly on V_{NMO} since the phase velocity depends linearly on V_{NMO}. From equations (5.8) and (5.9) it then follows that *the traveltime t becomes independent of V_{NMO} and that the associated offset x depends linearly on V_{NMO}*. Also, t_0 is a simple scaling factor for both the traveltime t and the associated offset x. In a single horizontal VTI layer, this means that *for fixed η and t_0, the moveout curve for different values of V_{NMO} can be calculated by simple horizontal stretching and squeezing along the offset axis* (see Figure 5.2). This important observation is a straightforward consequence of the negligible influence of V_{S0} on qP-wave traveltimes in TI media and the fact that the kinematics of qP-waves in homogeneous VTI media depend mainly on V_{NMO} and η (i.e., we can choose $\delta = 0$). To make explicit the independence of the traveltimes of V_{NMO}, we rewrite equation (5.8) as

$$t = \frac{t_0}{v|_{V_{NMO}=1}\cos\psi}\,, \tag{5.10}$$

where $v|_{V_{NMO}=1}$ denotes the group velocity calculated for $V_{NMO} = 1$ km/s.

Figure 5.1a (as well as b-d) shows that the contoured maximum traveltime differences between ray-traced traveltimes and traveltimes calculated using the nonhyperbolic moveout equation are independent of V_{NMO}. This is now easily understood in light of the previous observation that the nonhyperbolic moveout for fixed η (and t_0) but different values of V_{NMO}, are simply horizontally stretched or squeezed versions of each other. Note that we used $V_{S0} = 0$ and $\delta = 0$ for the ray tracing to determine the traveltime differences shown in Figure 5.1.

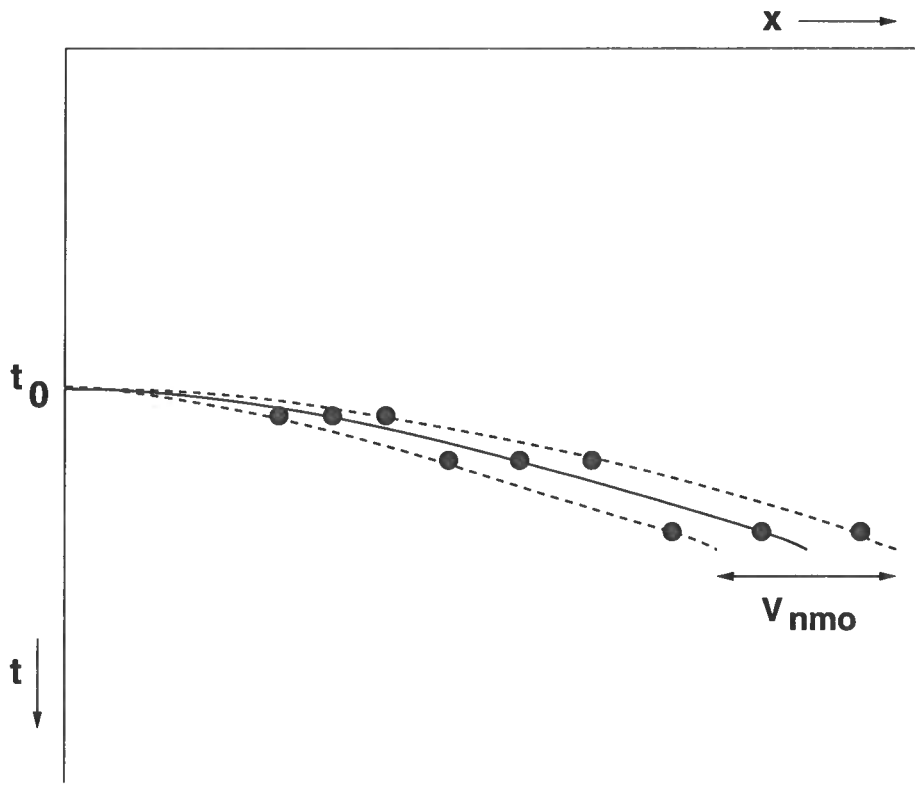


Figure 5.2. Under the customary assumptions that traveltimes of qP-waves in VTI media depend mainly on η and V_{NMO}, and that the shear-wave velocity along the symmetry axis (V_{S0}) has negligible influence on the traveltimes of qP-waves in such media, the nonhyperbolic moveout curve for fixed η and t_0, but varying V_{NMO}, can be calculated by simple horizontal stretching (or squeezing) along the offset axis.

Rewriting equation (5.1) gives

$$t^2(x) = t_0^2 + (x/V_{NMO})^2 - \frac{2\eta\,(x/V_{NMO})^4}{\left[t_0^2 + (1+2\eta)\,(x/V_{NMO})^2\right]}\,. \tag{5.11}$$

Using the observation that the offsets x are linear in V_{NMO}, i.e., using equation (5.9), it follows that

$$t^2(k) = t_0^2 \left\{1 + \frac{k^2}{4} - \frac{\eta k^4}{2\,(4 + [1+2\eta]\,k^2)}\right\}\,, \tag{5.12}$$

where we used $k/2 = \tan\psi$, with k the ODR. This equation shows explicitly that, under the above-mentioned approximations, the traveltimes for a given ODR k (or group angle

ψ) are independent of V_{NMO}. Note that using $k/2 = \tan\psi$ in equation (5.9) gives

$$k = \frac{2x}{t_0 V_{NMO}} \,. \tag{5.13}$$

Strictly, k is a true ODR only if $\delta = 0$. For convenience, we refer to k as the ODR throughout the remainder of this paper. Under the above-mentioned two approximations, the traveltimes $t(k)$ are independent of V_{NMO}.

5.5 Nonhyperbolic moveout analysis using rational interpolation

Fitting function values at various points (that are not necessarily distinct) using a rational function is usually referred to as *multipoint Padé approximation* (Baker & Graves-Morris, 1981). Such approximation is also referred to as *N-point Padé approximation*, *Newton Padé approximation*, or *rational interpolation*, depending on the context. Since in this paper, we do not use coincident interpolation points (often referred to as *confluent interpolation points*), we prefer to use the term *rational interpolation*.

A rational approximation to a function $T(x)$ is written as

$$T(x) \approx \frac{N_L(x)}{D_M(x)} \,, \tag{5.14}$$

where $N_L(x)$ is a polynomial of maximum order L, and $D_M(x)$ a polynomial of maximum order M. We denote such an approximation as $[L/M]$ and use the normalization $N_L(0) = T(0)$ and $D_M(0) = 1.0$, after (Baker, 1975, pp. 5-6). Given $n = L+M$ function values T_i at points x_i, with $i = 1, \ldots, n$, we arrive at a linear system of n equations with n unknowns, the coefficients of the polynomials. Once the coefficients are found, the resulting $[L/M]$ approximant can be used to find the function values $T(x)$ at values of x different from the interpolation points x_i. The solution to this system for the [2/2] rational approximation is given in Appendix M.

Here we use rational interpolation to approximate nonhyperbolic moveout in a horizontal transversely isotropic homogeneous layer with a vertical symmetry axis. Rewriting equation (5.1) in the form of a rational approximation using the definition of (Baker, 1975, pp. 5-6) gives

$$t^2(x) = t_0^2 \left(\frac{4 + 2\,(1+\eta)\,k^2 + 2k^4}{4 + (1+2\eta)\,k^2} \right) \,. \tag{5.15}$$

This expression reveals that the nonhyperbolic moveout equation can be viewed as a [2/1] rational approximation for t^2 as a function of k^2 (or x^2). Therefore, we could try a [2/1] rational interpolation to approximate nonhyperbolic moveout. In this paper, we choose a [2/2] rational interpolation in an attempt to gain extra accuracy.

To calculate the four unknown coefficients for the [2/2] rational interpolation using equations (M.2)-(M.5), we need four traveltimes and four associated offsets. Let t_i ($i = 1, \ldots, 4$) be the traveltime for a fixed ODR k_i (or group angle ψ_i), and let the associated offset be x_i. From the previous section, we know that the traveltimes t_i are independent of

V_{NMO} and depend on t_0 through a simple scaling only. This means that the traveltimes t_i needed for the rational interpolation can be calculated from a small subset of traveltimes calculated for a fixed reference value of t_0, denoted as t_0^{ref}, a range of η values, say from -0.2 to 1.0 in steps of 0.01. This subset can be precomputed and stored in a table. Hence, when the traveltimes t_i for a particular combination of t_0, V_{NMO}, and η are desired, a simple lookup in this table for the particular η combined with scaling with t_0/t_0^{ref} [cf. equation (5.10)], gives the traveltimes t_i needed for the rational interpolation. That is, the desired traveltime t_i is obtained from

$$t_i = t_i^{table} \left(\frac{t_0}{t_0^{ref}} \right) , \tag{5.16}$$

where t_i^{table} is the value of t_i obtained from the table (evaluated for $t_0 = t_0^{ref}$) for ODR k_i and the desired value of η. Evaluating the table for $t_0^{ref} = 1$ s allows the calculation of t_i to be done by scaling of t_i^{table} with t_0 only. Using the given values of V_{NMO} and t_0, the corresponding offsets x_i are then found simply from

$$x_i = \frac{V_{NMO}\, t_0\, k_i}{2} . \tag{5.17}$$

It is important to mention that the coefficients d_1, d_2, n_1, and n_2 (see Appendix M), needed for the rational interpolation for a given value of t_0, are calculated from traveltimes t_i and accompanying offsets x_i that depend only on η and V_{NMO}. Hence, even though we use a higher order rational approximation than the nonhyperbolic moveout equation from Alkhalifah and Tsvankin (i.e. [2/2] compared to [2/1]), the number of degrees of freedom are the same; i.e., both methods aim to resolve η and V_{NMO}, and thus resolve the same number of variables (two). This remains true even if we would use higher order rational interpolation compared to the [2/2] rational interpolation we use throughout this paper.

The traveltimes and offsets obtained using the method outlined above, allow us to perform velocity-analysis in VTI media using [2/2] rational interpolation. The efficiency of this approach is comparable to that of current velocity analysis using nonhyperbolic moveout equation (5.1), since the small subset of traveltimes for $t_0 = 1$ s and a range of η values, is precomputed once and stored in a table. In other words, the nonhyperbolic moveout equation is replaced simply with the rational interpolation formula, and the needed traveltimes are read from the precomputed table.

To precompute the table of traveltimes, a standard anisotropic ray-tracing algorithm can be used. Here, to calculate the traveltimes, we first solve

$$\frac{k_i}{2} = \frac{\tan\theta_i + \left. \frac{1}{V_i} \frac{dV}{d\theta} \right|_{\theta_i}}{1 - \left. \frac{\tan\theta_i}{V_i} \frac{dV}{d\theta} \right|_{\theta_i}} , \tag{5.18}$$

for θ_i numerically, using the Matlab function 'fsolve' which uses an interior-reflective Newton method to solve nonlinear equations (Coleman & Li, 1994, 1996). To obtain equation (5.18)

we used $\tan\psi_i = k_i/2$. In equation (5.18), V_i is the phase velocity associated with ODR k_i. Note that since we are ultimately interested only in traveltimes, we can use $\delta = 0$ and hence $\epsilon = \eta$. The value of $V_{P0} = V_{NMO}$ is irrelevant, since the term $\frac{1}{V_i}\frac{dV}{d\theta}\Big|_{\theta_i}$ is independent of V_{NMO} as explained earlier. Once θ_i is found, the traveltime t_i^{table} is then found through calculation of the group velocity $v_i|_{V_{NMO}=1}$ using equations (5.7) and (5.6) with $V_{NMO} = 1$ km/s, and subsequent use of this velocity in

$$t_i^{table} = \frac{t_0^{ref}}{v_i|_{V_{NMO}=1}\cos\left(\tan^{-1}\frac{k_i}{2}\right)}, \tag{5.19}$$

[cf. equation (5.10)]. We found that using the group angle as an initial guess for the phase angle generally worked well. We did not investigate different methods to solve for the phase angles θ_i. Using the 'fsolve' function in Matlab, the calculation of a table with four traveltimes for about 100 values of η takes on the order of one minute on a modern PC.

For hard rocks, setting $V_{S0} = 0$ is not as good an approximation as for softer rocks (Tsvankin, personal communication). For such geology (say carbonate reservoirs) it is better to use an 'intelligent estimate' of the V_p/V_s ratio. Note that in this case the traveltimes $t(k)$ are still independent of V_{NMO} because f is simply equal to some appropriately chosen constant value. Therefore, the moveout curves for fixed values of η and t_0 but different values of V_{NMO} are again horizontally stretched or squeezed versions of each other.

5.6 Accuracy comparison between rational interpolation and the nonhyperbolic moveout equation

Figure 5.1c shows contours of the maximum absolute percentage traveltime difference between ray-traced traveltimes for a single horizontal VTI layer and traveltimes resulting from the [2/2] rational interpolation. Here $t_0 = 1$ s, and the maximum ODR is two. The ODR values used for the rational interpolation are $k_1 = \frac{1}{2}$, $k_2 = 1$, $k_3 = \frac{3}{2}$, and $k_4 = 2$. Here the maximum traveltime errors are of the order $10^{-3}\%$ of t_0 (or 0.01 ms in this case), which is two-to-three orders of magnitude more accurate than that using the nonhyperbolic moveout equation. With a dominant frequency of 50 Hz, the traveltime errors are about 0.05% of the dominant period, and are hence negligible. Figure 5.1d shows the same contours, but now for a maximum ODR of four (the ODR values used for the rational interpolation are $k_1 = 1$, $k_2 = 2$, $k_3 = 3$, and $k_4 = 4$). The traveltime errors are now somewhat larger and of the order of $10^{-2}\%$ of t_0 (or 0.1 ms in this case), but still one-to-two orders of magnitude smaller than those for the nonhyperbolic moveout equation. Again, compared to a dominant period of 20 ms, these errors are negligible (0.5% of the dominant period). Hence, rational interpolation achieves a significant improvement in accuracy up to large ODR, and is highly accurate for all models of practical interest.

Figure 5.3 shows the residual moveout as a function of ODR for the nonhyperbolic equation (dashed) and for the [2/2] rational interpolation (solid). Here we used $V_{NMO} = 2.4$

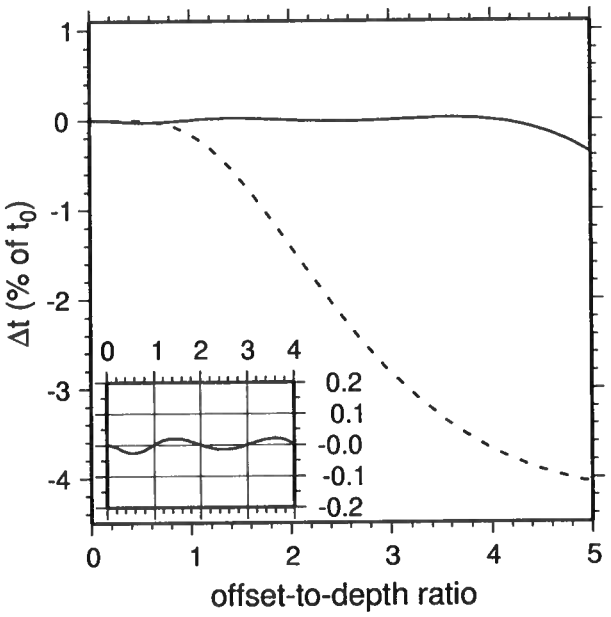


Figure 5.3. Residual moveout as a function of ODR using the nonhyperbolic moveout equation (dashed) and [2/2] rational interpolation (solid). In both cases the model parameters used are $\eta = 0.3$ and $V_{NMO} = 2.4$ km/s (indicated by the star in Figures 5.1a-c). The inset shows the residual moveout for the [2/2] rational interpolation on a larger scale (with axes as in the main figure). At the interpolated ODR values 1, 2, 3, and 4, the residual moveout is identical zero.

km/s [2], $\eta = 0.3$ (indicated by the star in Figures 5.1b and d), and $t_0 = 1$s. The inset shows the residual moveout for the [2/2] rational interpolation at a larger scale. The residual moveout is strictly zero at the specified ODR values of $k_1 = 1$, $k_2 = 2$, $k_3 = 3$, and $k_4 = 4$. We note that extrapolation beyond k_4 can result in large errors in traveltimes because the [2/2] rational approximation starts oscillating with large amplitudes. The onset of such oscillation can be seen in Figure 5.3 for ODR values beyond $k_4 = 4$. This is caused by fact that in the limit of $x \rightarrow \infty$, a [2/2] rational approximation becomes a constant. Therefore, for our [2/2] rational interpolation, k_4 should be chosen such that it exceeds the largest offset in the data.

Comparing Figures 5.1c and 5.1d, it seems that use of larger intervals Δk between the k_i causes the maximum traveltime errors to increase. To analyze the influence of Δk, and to see what maximum ODR can be reached with [2/2] rational interpolation with considerable accuracy, Figure 5.4 shows the maximum absolute percentage traveltime error as a function of Δk. The top horizontal scale of the figure shows the maximum ODR. Here we use $V_{NMO} = 2.4$ km/s and $t_0 = 1$ s, and plot the traveltime errors for two different values of η,

[2]Because of our observation that the traveltime errors do not depend on V_{NMO}, the actual value of V_{NMO} is of course irrelevant.

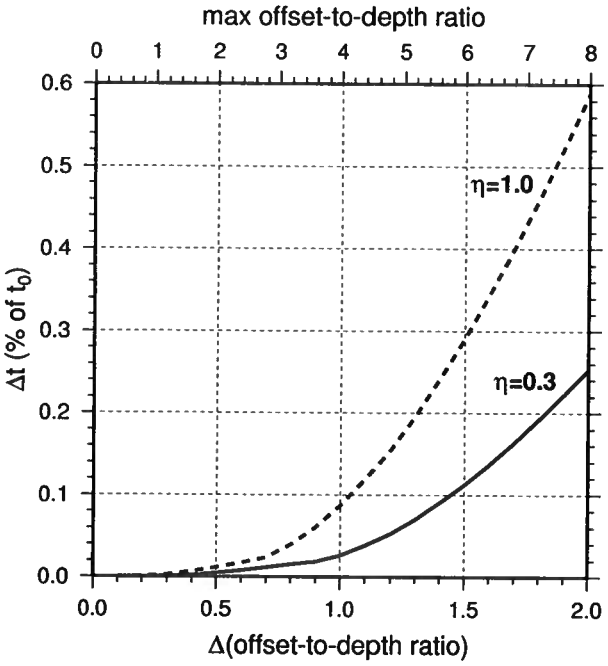


Figure 5.4. Accuracy of the [2/2] rational interpolation as a function of the separation in ODR of the interpolation traveltimes for $\eta = 0.3$ and $\eta = 1.0$. In both cases $V_{NMO} = 2.4$ km/s. Even if the separation in ODR is two, i.e., the maximum ODR is 8, the percentage errors in traveltime are of the order $10^{-1}\%$.

i.e., $\eta = 0.3$ (solid) and $\eta = 1.0$ (dashed). Note again that the value of V_{NMO} is irrelevant; errors do not depend on V_{NMO}, as explained previously. The high η value for the dashed curve can be taken as a worst-case scenario with respect to accuracy. For practical values of η, the maximum error in traveltime for ODR values up to 8 (achieved with $k_1 = 2$, $k_2 = 4$, $k_3 = 6$, and $k_4 = 4$) is of the order of a tenth of a percent of t_0. Therefore, for almost all practical cases, our [2/2] rational interpolation provides an accurate nonhyperbolic moveout approximation. If needed, higher-order rational interpolation, combined with $\Delta k = 1$, would obtain extra accuracy.

Figure 5.5 shows semblance scans (at fixed t_0) as a function of V_{NMO} and the horizontal velocity $V_{hor} = V_{NMO}\sqrt{1+2\eta}$, calculated using the nonhyperbolic moveout equation (a, c, and e) and the [2/2] rational interpolation (b, d, and f). In each column of subfigures the true model parameters (indicated by the star) vary, and are, respectively, $V_{NMO} = 2000$ m/s and $V_{hor} = 2300$ m/s (or $\eta = 0.16$) for a and b, $V_{NMO} = 2892$ m/s and $V_{hor} = 3745$ m/s (or $\eta = 0.34$) for c and d, and $V_{NMO} = 2464$ m/s and $V_{hor} = 3880$ m/s (or $\eta = 0.74$) for e and f; the semblance maxima are indicated by the squares. Note that the synthetic gathers were generated using ray-traced traveltimes, where the ray tracing was done using the true values of δ, i.e., $\delta = 0$ for a and b, $\delta = -0.05$ for c and d, and $\delta = -0.22$ for e and f, while we used $V_{S0} = 0$ km/s for all models.

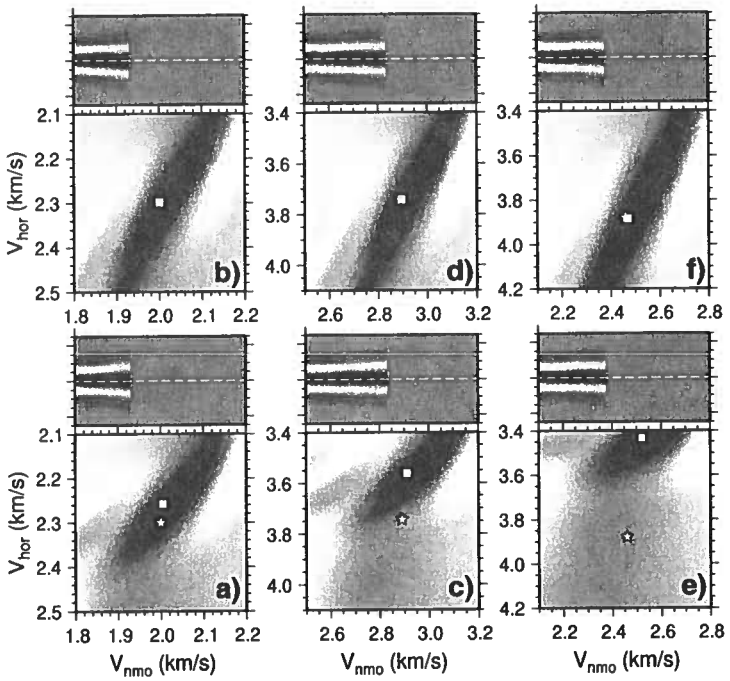


Figure 5.5. Semblance scans and moveout-corrected gathers, where the moveout correction was done with the parameters related to the maximum semblance values (indicated by the square). The model for each subfigure consists of a single horizontal VTI layer, and the maximum ODR is two. For all models, $t_0 = 1$ s. The true model parameters are indicated by the star, and are: $V_{NMO} = 2.0$ km/s and $V_{hor} = 2.3$ km/s (i.e., $\eta = 0.16$) in subfigures a and b; $V_{NMO} = 2.892$ km/s and $V_{hor} = 3.745$ km/s (i.e., $\eta = 0.34$) in c and d (shale under zero confining pressure); and $V_{NMO} = 2.46$ km/s and $V_{hor} = 3.88$ km/s (i.e., $\eta = 0.74$), in e and f (Green River Shale). In subfigures a, c, and e, the nonhyperbolic moveout equation was used, while in b, d, and f, we used the [2/2] rational interpolation.

The model parameters for Figures 5.5a and b correspond to the model parameters used in Figures 1 and 2 of Grechka & Tsvankin (1998). [These parameters were originally chosen because such values of η were observed on field data (Alkhalifah *et al.*, 1996)]. Since these model parameters are identical to those in Figures 1 (and 2) of Grechka & Tsvankin (1998), our Figure 5.5a is the semblance scan equivalent of Figure 1 in their paper. For Figures 5.5c and d the model parameters correspond to a shale under zero confining pressure, and for Figures 5.5e and f the parameters correspond to Green River shale [see Table 1 in Thomsen (1986) for these two cases]. For all subfigures (i.e., a - f) we used $t_0 = 1$ s and a maximum ODR of two. For the rational interpolation we used $k_1 = \frac{1}{2}$, $k_2 = 1$, $k_3 = \frac{3}{2}$, and $k_4 = 2$.

Figures 5.5a, c, and e show that the nonhyperbolic moveout equation (5.1) obtains high semblance values, but for the wrong values of V_{hor}; V_{NMO} seems largely unbiased. This means that the associated common midpoint (CMP) gathers are well flattened using a

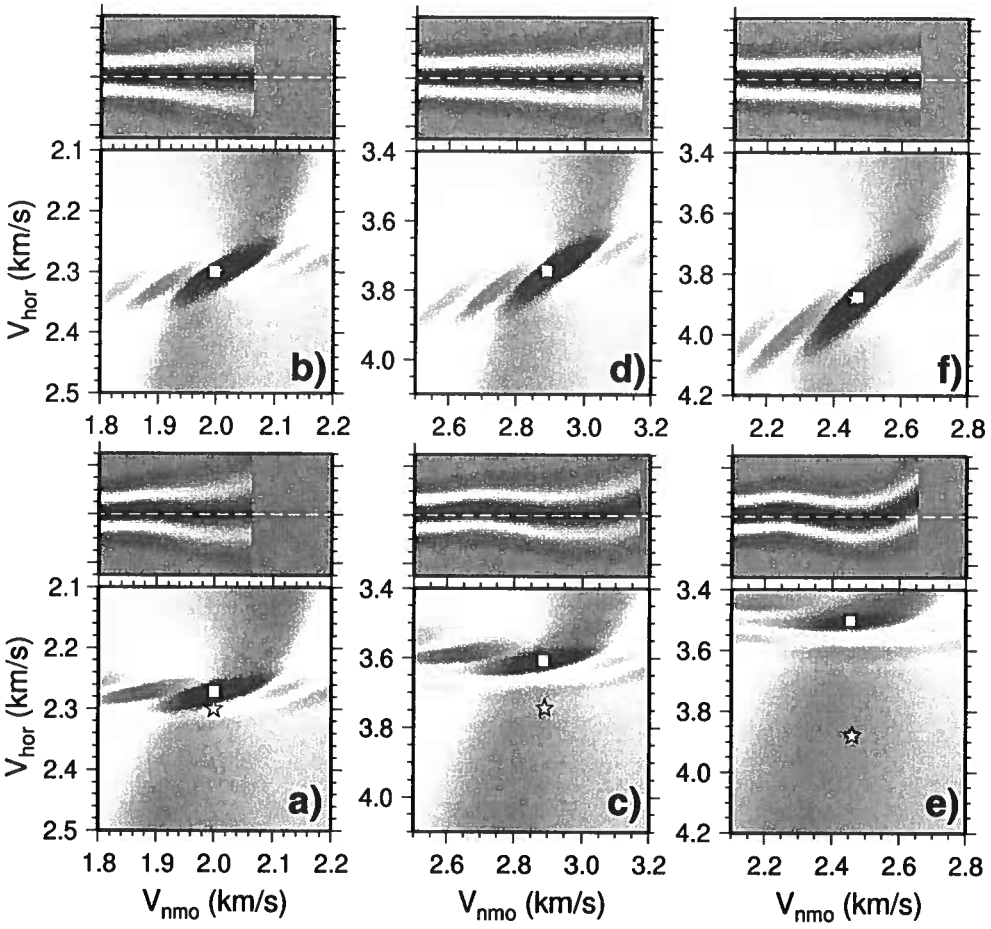


Figure 5.6. Same as Figure 5.5, except that the maximum ODR is now four.

biased estimate of V_{hor}, hence introducing a bias in the estimated value of η. The associated moveout corrected gathers for all three models are shown above the semblance scans. Note that for all three models, the semblance maximum (indicated by a square) for the nonhyperbolic moveout equation indicates a value of η smaller than the true value (a smaller difference between V_{hor} and V_{NMO} than the difference between their true respective values). Hence, this method seems to underestimate η. The nonhyperbolic moveout equation obtains accurate η values only when the true η values are small (≤ 0.1). Straightforward application of the [2/2] rational interpolation method provides, for all models, maximum semblances that coincide with the true model parameters, and accurately flattens gathers without any residual moveout (see Figures 5.5b, d, and f).

Figure 5.6 shows the same semblance scans and moveout corrected gathers, for the same models and methods as in Figure 5.5, except that here the maximum ODR is four. For the rational interpolation we used $k_1 = 1$, $k_2 = 2$, $k_3 = 3$, and $k_4 = 4$. Notice that for all models and both methods, the peak of the semblance scans is much better defined than

in Figure 5.5; that is, the large semblance values span a much smaller range of V_{hor} (and thus η) values, and a somewhat smaller range of V_{NMO} values. Indeed the higher resolution for η for larger offset ranges was mentioned by Alkhalifah (1997), Grechka & Tsvankin (1998), and Wookey *et al.* (2002). Aside from the improved resolution in η, Figures 5.6a, c, and e, show again the underestimation of η when the nonhyperbolic moveout equation is used, just as in Figures 5.5a, c, and e. The maximum semblance, however, now deteriorates with increasing levels of anellipticity, and the associated moveout-corrected CMP gathers clearly indicate a distortion in the residual moveout, even for the model with $\eta = 0.16$ (see Figure 5.6a), a value that is commonly observed on field data [e.g. Alkhalifah *et al.* (1996)]. The rational interpolation method (Figures 5.6b, d, and f) results in no residual moveout *combined with* unbiased estimates of η and V_{NMO}. This is a direct consequence of the high accuracy of the rational interpolation method shown in Figure 5.1d for this ODR. If accuracy is desired for even larger offsets, higher-order rational interpolation can be used. Alternatively, [2/2] rational interpolation can be used with an increased interval between the ODR k_i (say $k_1 = 2$, $k_2 = 4$, $k_3 = 6$, $k_4 = 8$), as Figure 5.4 indicated that with the [2/2] rational interpolation we can achieve reasonable accuracy up to ODR of 8.

5.7 Accuracy comparison with the nonhyperbolic moveout approximation of Fomel

Recently, Fomel (2004) proposed a shifted-hyperbola approximation for the group velocity and converted this into the following moveout equation for a single homogeneous horizontal VTI layer,

$$t^2(x) = \frac{3+4\eta}{4\,(1+\eta)} H(x) + \frac{1}{4\,(1+\eta)} \sqrt{H^2(x) + 16\eta\,(1+\eta)\,\frac{t_0^2 x^2}{(1+2\eta)\,V_{NMO}^2}}\,, \tag{5.20}$$

where $H(x)$ represents the hyperbolic part,

$$H(x) = t_0^2 + \frac{x^2}{(1+2\eta)\,V_{NMO}^2}\,. \tag{5.21}$$

Fomel (2004) showed that this approximation is significantly more accurate than the nonhyperbolic moveout equation (5.1) of Alkhalifah & Tsvankin (1995).

The rational interpolation approach that makes use of the fact that the influence of V_{NMO} on the moveout is limited to a stretch along the offset axis, is based on the approximations that (1) the influence of V_{S0} on qP-wave traveltimes in TI media is negligible, and that (2) the kinematics of qP-waves in homogeneous VTI media depend mainly on V_{NMO} and η, i.e., we can set $\delta = 0$. Figure 5.7 shows the traveltime difference between ray-traced traveltimes for a single horizontal VTI layer with Greenhorn-shale anisotropy (Jones & Wang, 1981) and the nonhyperbolic moveout equation (5.1) form Alkhalifah and Tsvankin (dotted), the shifted hyperbola equation (5.20) from Fomel (dashed), and traveltimes resulting from ray tracing with $V_{S0} = 0$ km/s and $\delta = 0$ but the same values of V_{NMO}

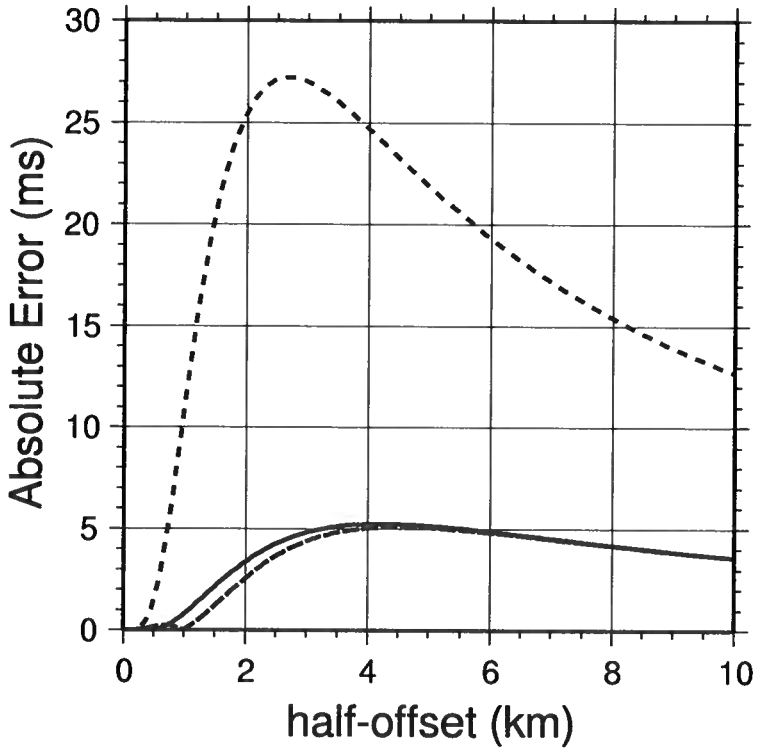


Figure 5.7. Comparison of accuracy between the nonhyperbolic moveout equation without the correction factor (dotted), the shifted hyperbola method of Fomel (2004) (dashed), and our rational interpolation approach (solid), for Greenhorn-shale anisotropy ($\eta = 0.34$ and $V_{NMO} = 2935$ m/s). The absolute traveltime errors are calculated with respect to ray-traced traveltimes. The model parameters are: $V_{P0} = 3094$ m/s, $V_{S0} = 1510$ m/s, $\delta = -0.05$, and $\epsilon = 0.256$ (Greenhorn shale).

and η (solid). The ansitropy parameters are $V_{P0} = 3094$ m/s, $V_{S0} = 1510$ m/s, $\delta = -0.05$, and $\epsilon = 0.256$ (i.e., $V_{NMO} = 2935$ m/s and $\eta = 0.34$), and $t_0 = 646.5$ ms. Except for the solid line, Figure 5.7 reproduces Figure 7 of Fomel (2004). Note that the accuracy of the shifted hyperbola approximation is marginally better than the ray-traced traveltimes using $V_{S0} = 0$ km/s and $\delta = 0$. The maximum difference between these two traveltime approximations is about 1 ms (i.e., 0.15% of t_0) for this particular model. From a practical point of view, the two approximations are therefore identical. Even though we have not calculated the solid line with a rational interpolation, we can approximate the solid line to almost arbitrary precision with rational interpolation (i.e., with $[M/N]$ rational interpolation where $M, N > 2$). For ODR up to four, we showed this in Figure 5.1d using a [2/2] rational interpolation. Since the rational interpolation method that we propose uses traveltimes calculated with $V_{S0} = 0$ km/s and $\delta = 0$, we conclude that the accuracy of our method is basically identical to that of the shifted hyperbola approximation of Fomel

(2004). Comparison of semblance scans and moveout corrections calculated using equation (5.20) and the rational interpolation method (not shown here), for the models and offsets studied in Figures 5.5 and 5.6, showed no difference between both methods.

5.8 Application to horizontally layered VTI media

Up to this point, we have treated only a single horizontal VTI layer. In this section we test the applicability of the rational interpolation method to a horizontally layered VTI medium, and compare its accuracy to that of the nonhyperbolic moveout equation. Based on the work of Tsvankin & Thomsen (1994), Grechka & Tsvankin (1998) showed that the nonhyperbolic moveout equation (5.1) remains valid in horizontally layered media provided the parameters η and V_{NMO} are replaced by 'effective' values that are some average over the vertically heterogeneous overburden. Also, they rewrite the Dix-type inversion procedure, originally introduced by Tsvankin & Thomsen (1994), in terms of η and V_{NMO}. In that method, the effective values of η and V_{NMO}, obtained from applying the single-layer equation (5.1) to data from a layered medium, are used to estimate the interval values of η and V_{NMO}.

We compare the effective values of η and V_{NMO} obtained using the nonhyperbolic moveout equation and the rational interpolation method, through application of both methods to a horizontally layered medium with the following parameters:

- layer 1 : $h = 1$ km (h is the depth of the bottom of the layer), $V_{NMO} = 2098$ m/s, $V_{hor} = 2098$ m/s, $\eta = 0$, ($V_{P0} = 2000$ m/s, $\epsilon = 0.05$, $\delta = 0.05$),
- layer 2 : $h = 2$ km, $V_{NMO} = 2000$ m/s, $V_{hor} = 2300$ m/s, $\eta = 0.16$, ($V_{P0} = 2000$ m/s, $\epsilon = 0.16$, $\delta = 0$),
- layer 3 : $h = 3$ km, $V_{NMO} = 2892$ m/s, $V_{hor} = 3745$ m/s, $\eta = 0.34$, ($V_{P0} = 3048$ m/s, $\epsilon = 0.255$, $\delta = -0.05$),
- layer 4 : $h = 4$ km, $V_{NMO} = 2460$ m/s, $V_{hor} = 3880$ m/s, $\eta = 0.74$, ($V_{P0} = 3292$ m/s, $\epsilon = 0.195$, $\delta = -0.22$).

For all models we used $V_{S0} = 0$ km/s. Note that here the first layer is elliptically anisotropic, and the second through fourth layers have the same model parameters as those in the models studied in the first through third row of Figures 5.5 and 5.6, respectively.

Table 5.1 shows the results from both methods for a maximum ODR of two (top) and four (bottom). Since both methods are for a single-layer VTI medium only, and because we apply these methods to layered VTI media, the η, V_{NMO}, and V_{hor} values in Table 5.1 are all 'effective' values. The results closely resemble the results obtained from the single layer numerical tests shown in Figures 5.5 and 5.6. For maximum ODR=2, the nonhyperbolic moveout equation gives consistently lower estimates of η than the rational interpolation method, and the difference increases as we increase the level of anellipticity, i.e., as we go deeper. At the same time, the semblance values (indicated with s in Table 5.1) are high

Table 5.1. Comparison of obtained semblance values (s), V_{NMO}, V_{hor}, and η, for both methods (NHME = nonhyperbolic moveout equation, and RI is rational interpolation), when applied to a layered medium and maximum ODR of two (top) and four (bottom). The model consists of four horizontal homogeneous VTI layers.The parameters for each layer are given in the main text.

ODR = 2

	NHME				**RI**			
layer	s	V_{NMO}	V_{hor}	η	s	V_{NMO}	V_{hor}	η
1	1.00	2097	2101	0.00	1.00	2099	2101	0.00
2	0.98	2058	2179	0.06	0.98	2058	2188	0.07
3	1.00	2271	2839	0.28	1.00	2252	3022	0.40
4	0.99	2313	3012	0.35	1.00	2282	3285	0.54

ODR = 4

	NHME				**RI**			
layer	s	V_{NMO}	V_{hor}	η	s	V_{NMO}	V_{hor}	η
1	1.00	2094	2099	0.00	1.00	2096	2099	0.00
2	0.95	2029	2207	0.09	0.96	2035	2218	0.09
3	0.57	2238	2894	0.34	0.84	2189	3158	0.54
4	0.57	2320	3038	0.36	0.87	2241	3391	0.64

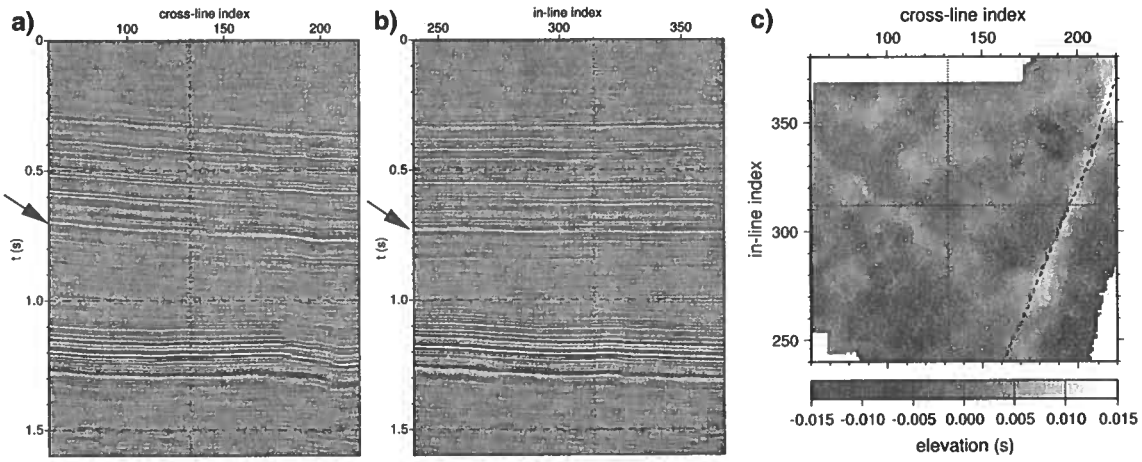


Figure 5.8. Field data used to test the rational interpolation method. Inline (a), crossline (b), and plan view of the residual topography of the event used for testing (indicated by the arrows in a and b) after removal of the regional dip (c). The dotted lines in c denote the locations of the inline and crossline sections shown in a and b. The dashed line in c denotes the location of a regional fault.

for all layers using both methods, indicating high quality moveout correction. This is analogous to the behaviour shown in the single layer numerical test shown in Figure 5.5; i.e., both methods provide accurate moveout correction, but the nonhyperbolic moveout equation does so with a biased (too low) value of η. For maximum ODR=4, the nonhyperbolic moveout equation results in decreasing semblance with increasing level of anellipticity, (i.e. as we go deeper), indicating a lack of ability to accurately moveout-correct the data. That the rational interpolation method has substantially larger semblance values than the nonhyperbolic moveout equation, indicates that this method is able to flatten the gathers best, with the least residual moveout. Again this is analogous to the results from the single-layer case (Figure 5.6). Hence, it seems that the rational interpolation method, which is based on a single horizontal VTI layer, is suitable for application to a horizontally layered medium, at least up to maximum ODR=4 and high levels of anellipticity.

5.9 Field data example

Figures 5.8a and b show inline and crossline stacked time sections from a land dataset that contains a reflector at about 750-ms two-way traveltime (indicated by the arrows), illuminated with ODR ranging to larger than four. We focus attention on this event throughout the remainder of this section. The geology consists of relatively flat (dip less than two degrees), predominantly shale layers, such that a layered VTI model seems, at first sight, appropriate for these data. Although the structure on the horizon of interest is limited to within ± 20 ms of a best-fit planar dip, a subtle NNE-SSW structural trend, associated with

deeper faulting (Figure 5.8c), exists. The dashed line indicates the location of a regional fault. The inline and crossline spacings are both 110 ft.

To test our method, we calculated semblance as a function of V_{NMO} and V_{hor} for the whole dataset over a 100-ms window centered on the event of interest. The offsets used in the analysis were limited to offsets with a maximum ODR of approximately four. Figure 5.9a and b show the semblance scans for both methods for a randomly selected CMP gather. The location of this gather is indicated by the intersection of the dotted lines in Figure 5.8c; the location is also indicated by the vertical dotted lines in Figures 5.8a and b. This gather was generated by collecting traces from a 3-by-3 super bin of adjacent CMPs, and subsequent offset-binning. The derived η, V_{NMO}, and maximum semblance values for both methods are: 0.34, 3100 m/s, 0.38 (nonhyperbolic moveout equation) and 0.33, 3250 m/s, 0.40 (rational interpolation). The inferred values of η are close to the η value used to generate the results shown in Figures 5.6c and d, where $ODR = 4$ also. Note the striking similarity between the change of shape of both the moveout corrected gathers and the semblance scans of the field data (Figures 5.9a-d) and the synthetic data (Figure 5.6c and d), as we change from the nonhyperbolic moveout equation to the rational interpolation method. For the semblance scans, the nonhyperbolic moveout equation method exhibits no clear evidence of the inherent trade-off relation between η and V_{NMO}, whereas the rational interpolation method does display this known trade-off. Also, the semblance peak is most clearly defined for the [2/2] rational interpolation method because of its higher accuracy at larger ODR. Note that the residual moveout for the nonhyperbolic moveout equation is substantial, whereas the rational interpolation method gives well corrected moveout.

Figures 5.10a and b show mapviews of the values of η obtained for the event of interest over the entire dataset, using, respectively, the (a) nonhyperbolic moveout equation and (b) the [2/2] rational interpolation method, for maximum ODR of two. The η values obtained using nonhyperbolic moveout equation, are generally smaller than the estimated η values from the rational interpolation method, just as in our synthetic data examples (cf. Figure 5.5). The same comparison for V_{NMO} (not shown) indicated that both methods gave similar estimates of V_{NMO}, just as in the synthetic examples. Figures 5.10c and d are as a and b, but here the maximum ODR is four. Notice, that the estimated η values are spatially less variable, for both methods. This can be understood in light of the improved resolution in η for larger ODR. For both $ODR = 2$ and $ODR = 4$, the [2/2] rational interpolation results in more spatially smooth and continuous values of η (and V_{NMO}) than the nonhyperbolic moveout equation. It is important to note that this spatial continuity was not imposed, but followed from a straightforward application of the rational interpolation method presented here. Since the seismic data show no indication of substantial lateral heterogeneity (cf. Figure 5.8), the relatively smooth spatial variation of η is consistent with the seismic data. This increases our confidence in the η values obtained using the rational interpolation method. In some parts the values of η seem to correlate somewhat with the geologic trend (cf. Figure 5.8c).

Figure 5.11a shows the normalized semblance difference between the nonhyperbolic moveout equation method and the rational interpolation method, for maximum ODR=2. The normalized semblance difference is the difference between the semblances from both

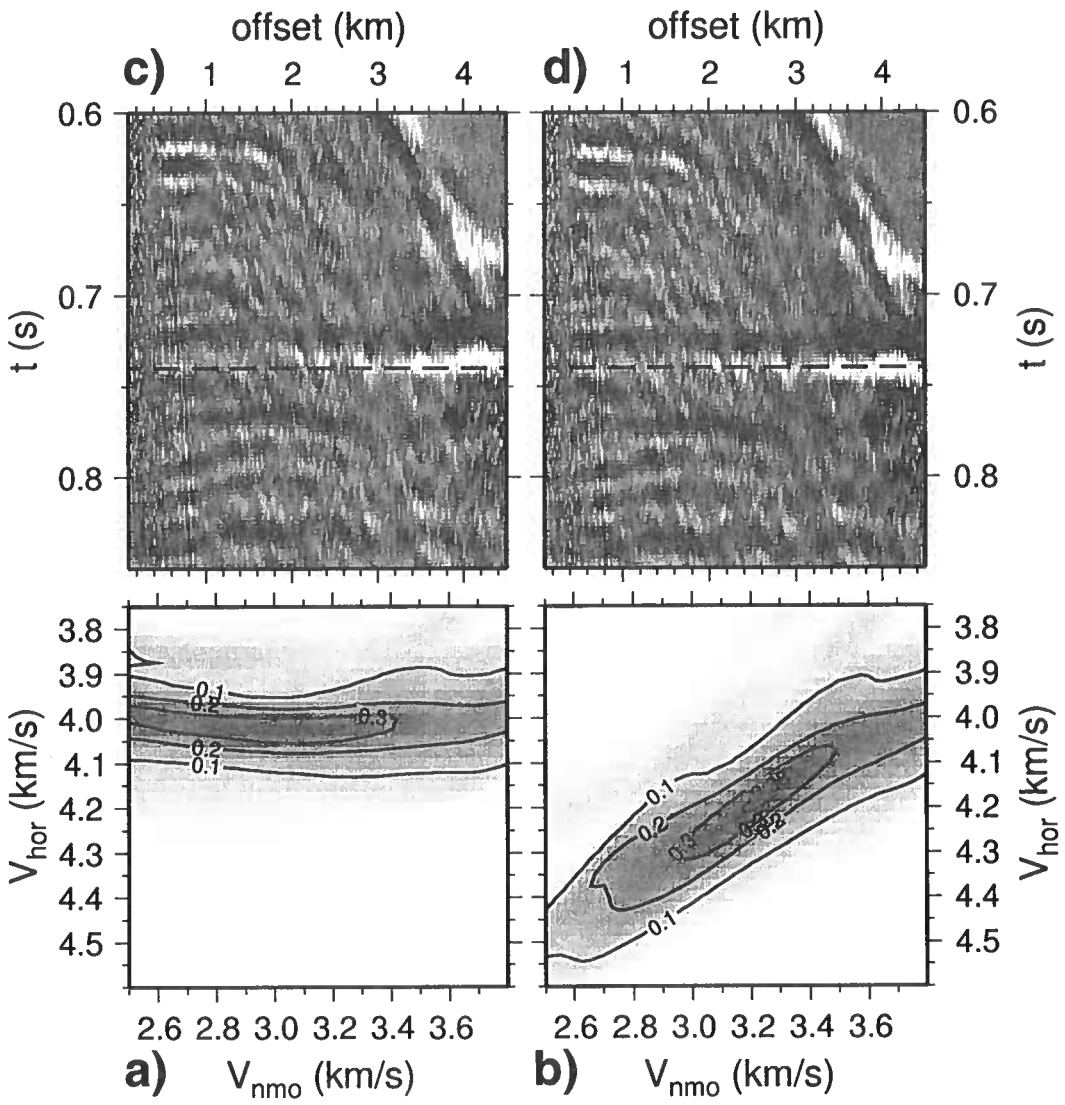


Figure 5.9. Semblance scans (a and b) for one CMP gather, located at the crossing of the horizontal lines in Figure 5.8c; its location is also indicated by the dotted lines in Figures 5.8a and b. The associated moveout corrected gathers are shown in subfigures c and d. The methods used for parameter estimation (i.e., to calculate the semblance scans) and moveout correction are, respectively, the nonhyperbolic moveout equation (a and c) and the [2/2] rational interpolation method (b and d). A 100-ms window centered around $t_0 = 750$ ms was used in the computation. The contours in a and b indicate the semblance values. Maximum semblance values for both methods are, respectively, 0.37 (a) and 0.40 (b), and the resulting estimates for η and V_{NMO} are, respectively, $\eta = 0.34$ and $V_{NMO} = 3100$ m/s (a), and $\eta = 0.33$ and $V_{NMO} = 3250$ m/s (b).

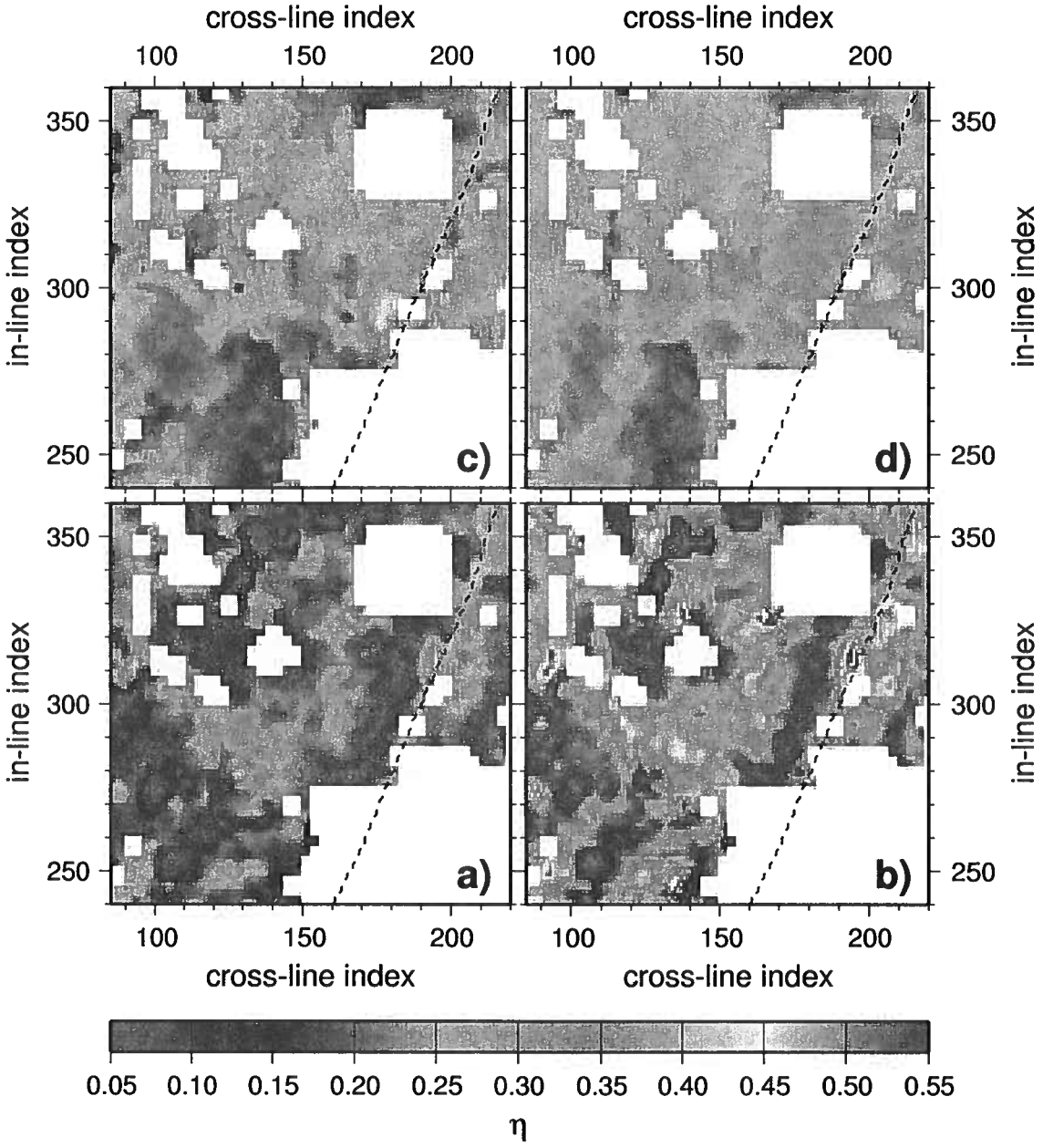


Figure 5.10. Map view of η derived from the event of interest in the field data with maximum offset of 2250 m, i.e., $ODR \approx 2$ (a and b), using semblance scans with the nonhyperbolic moveout equation (a) and the [2/2] rational interpolation method (b). Subfigures c and d are as a and b, except that the maximum offset is 4500 m, i.e., $ODR \approx 4$. Estimates of η at locations with poor offset distributions were set to white.

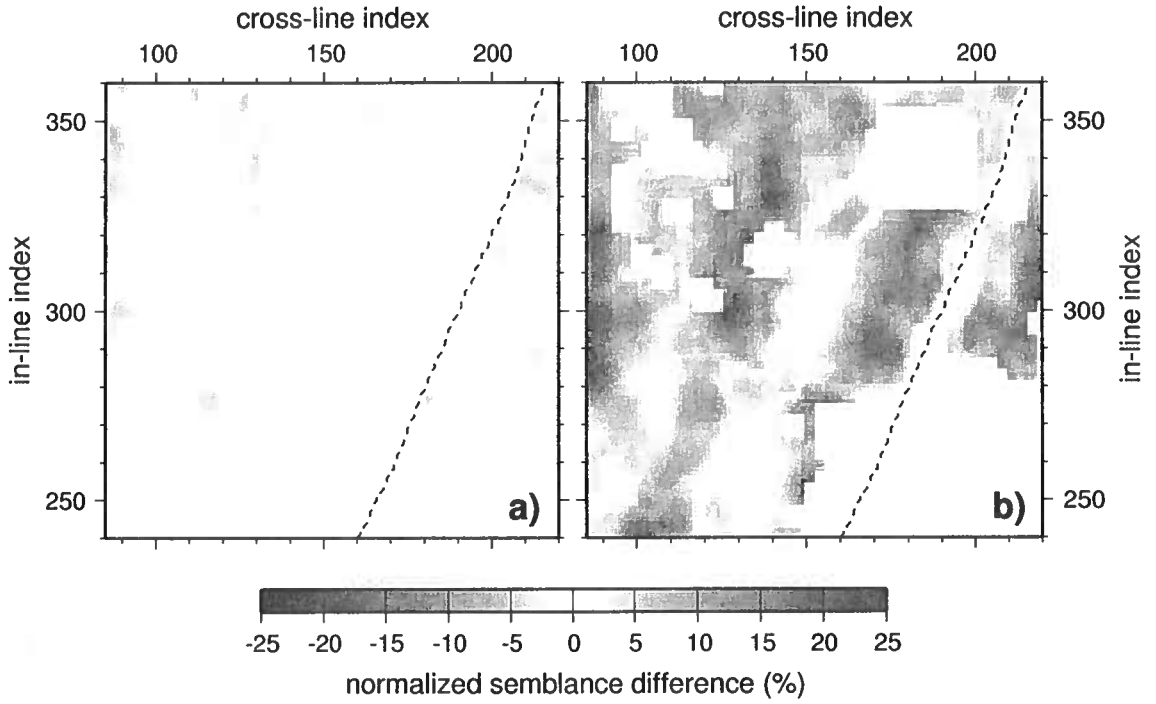


Figure 5.11. Normalized semblance difference between the nonhyperbolic moveout equation and the [2/2] rational interpolation method, for a maximum offset of 2250 m, i.e., $ODR \approx 2$ (a) and maximum offset of 4500 m, i.e., $ODR \approx 4$.

methods divided by the semblance for the nonhyperbolic moveout equation method. For these offsets, both methods result in similar semblances, indicating similar ability to flatten the gathers. This supports our findings from the numerical tests that both methods obtain comparable results in terms of moveout correction, for an ODR range up to two (cf. Figure 5.5). Figure 5.11b is as 5.11a, except that here the maximum ODR is four. For this range of ODR, the [2/2] rational interpolation has on average 10% higher semblance values than does the nonhyperbolic moveout equation method. This supports (for the whole field dataset) our findings from the synthetic results. Hence, for this field data example and for a maximum ODR larger than two, the [2/2] rational interpolation method provided generally improved moveout correction (i.e., improved semblance values), and more spatially continuous estimates of η.

5.10 Discussion

We have shown that for ODR values up to 8, the [2/2] rational interpolation results in an accuracy of $O(10^{-1})\%$, or less, of t_0 for most models of practical interest. If higher accuracy is desired, or accuracy up to larger ODR is needed (e.g., in near-surface seismic experiments), higher-order polynomials can be used in the rational approximation. The

added computation time for inclusion of several extra terms is negligible; hence the efficiency of the proposed method remains essentially the same.

As with any linear system of equations, the system used to calculate the coefficient n_1, n_2, d_1, and d_2, may be singular. In our case, this happens when the moveout is purely hyperbolic (see Appendix N). However, as explained by Stoer & Bulirsch (1993, p.61-62), this nonsolvability of the interpolation problem is a matter of degeneracy; hence, solvability can be restored by arbitrary small perturbations of the support points. For purely hyperbolic moveout we indeed observed that adding a small amount of numerical noise (on the order of 10^{-2} ms) to the support traveltimes, restored solvability. Figure 5.1, however, shows that for virtually all anisotropic models of practical interest such degeneracy does not occur, indicating that for our [2/2] rational interpolation we need not worry about such degeneracy. We have not tested for degeneracy using higher order rational interpolation.

Rational interpolation can be accompanied with the presence of (nearby)[3] unwanted poles on (or near) the interval of interpolation [e.g., Berrut & Mittelmann (1997, p.357) or Berrut & Mittelmann (2000)]. Once the coefficients n_1, n_2, d_1, and d_2 are determined, the presence of such poles can easily be verified by checking for zeroes in the denominator of the corresponding [2/2] rational approximation. Also, because in the absence of nearby poles in the complex plane the derivative of the traveltime curve should always be positive in a single horizontal VTI layer, the presence of such (nearby) poles can be checked for by checking for a change in sign of the derivative of the rational approximation $t(x)$. When a (nearby) pole is found, slight perturbation of the support points using noise injection should remove the pole. However, Figure 5.1 indicates that for virtually all anisotropic models of practical interest we did not observe the presence of such poles.

We mentioned in the above that the [2/2] rational interpolation converges to a constant in the limit of infinite offset, causing the interpolant to oscillate with large amplitudes beyond the maximum interpolation ODR k_4. This generally holds for $[M/N]$ rational interpolation where $M = N$. However, using for example $M = N + 1$, then the limit of the interpolant converges to a linear function. Even more so, setting the coefficient $n_{M-1} = 0$, the interpolant $T(X)$ then converges to the function $\frac{n_M}{d_{M-1}}X$, which has the right functional dependence for the moveout curve at infinite offset. This would possibly allow extrapolation beyond the maximum interpolation ODR k_i with smaller error. We have not tested any such approach. For this paper, the [2/2] rational interpolation suffices to exemplify the accuracy of the rational interpolation in the context of velocity analysis in VTI media.

The geometry of straight rays involved in velocity analysis for a horizontal homogeneous layer and the geometry of the rays associated with point scattering in a medium with constant velocity are identical. Therefore, the rational interpolation method proposed here is immediately applicable to the problem of post-stack and pre-stack time-migration in VTI media. Larger offsets in the context of moveout velocity analysis are the equivalents of steeper dips in time-migration. Therefore, the increased accuracy at larger ODR

[3]Here, nearby is meant in the complex plane; in practice nearby would likely be defined as having a distance (in the complex plane) smaller than the smallest distance between the support abscissas x_i and x_{i+1}.

provided by the rational interpolation, suggests improved accuracy when the rational interpolation scheme is used in imaging of steep reflectors in the context of pre and post-stack time-migration in VTI media. In this context, our observation that, for fixed η and t_0, the moveout curve for different values of V_{NMO} can be calculated by simple horizontal stretching and squeezing, implies that the pre and post-stack migration operators for different values of V_{NMO} are also found by such stretching and squeezing.

None of our synthetic tests included amplitude-versus-offset (AVO) variations, whereas the field data example clearly does (cf. Figure 5.9). Some additional synthetic tests (not shown), indicate that the rational interpolation method is less sensitive to such AVO variations than is the nonhyperbolic moveout equation method; this can be explained by the high accuracy of the rational interpolation method[4]. Note that the presence of AVO variations causes semblance based moveout correction to be biased towards the offsets with higher amplitudes. Such bias is especially noticeable when an inaccurate moveout approximation is used to moveout correct the data.

The average η values resulting from application of the rational interpolation method and the nonhyperbolic moveout equation method to the field data for a maximum ODR=4, are $\eta_{av} = 0.28$ and $\eta_{av} = 0.25$, respectively. For both methods, these values of η are high compared to values typically reported from nonhyperbolic moveout analysis [e.g. Toldi *et al.* (1999)]. Of course the relatively high η values are likely a result of the particular lithology of this area. However, most studies reporting estimates of η have been done on marine data, where a substantial waterlayer reduces the effective values of η, hence possibly introducing a bias in what are traditionally considered acceptable values of η. Of course, the underestimation of η using the nonhyperbolic moveout equation as pointed out in this work, only aggravates possible underestimation of acceptable values of η.

In the field data example, we treat the overburden of the reflection event of interest as a single horizontal homogeneous VTI layer. As a result, the estimated values of η and V_{NMO} are effective parameters. The geological significance of such effective quantities is difficult to establish, and an approach assuming a layered overburden and resulting interval estimates of η and V_{NMO} would remove this difficulty. Although we do not demonstrate such an approach in this paper, we believe that such an approach is feasible by applying rational interpolation in a layer-stripping fashion. We leave the verification of this idea to a future study. Meanwhile, the single-layer approach outlined here can be used to obtain more accurate estimates of the average (or effective) values of η and V_{NMO} in layered media. Using these values in current Dix-type averaging procedures could lead to better interval estimates of these parameters.

For horizontally layered media, a plane-wave decomposition, obtained through a $\tau - p$ transform, is the natural decomposition of the data. This fact was successfully used by Van der Baan & Kendall (2002) and Van der Baan (2004), to obtain interval estimates of η and V_{NMO} from moveout in the $\tau - p$ domain. Even though they successfully estimate interval values of η and V_{NMO}, the need to transform the data to the $\tau - p$ domain or to pick traveltimes in the $t - x$ domain, is a practical disadvantage. Therefore, there is room

[4]We did not include phase changes with offset in the tests concerning sensitivity to AVO variation.

to try to extend our $t-x$ based rational interpolation method to render interval estimates of η and V_{NMO}.

5.11 Conclusions

We have presented a rational interpolation approach to nonhyperbolic moveout analysis and correction of qP-waves in VTI media. The accuracy of the method was tested using both synthetic and field data and compared with that of the nonhyperbolic moveout equation of Alkhalifah and Tsvankin (Alkhalifah & Tsvankin, 1995), which is the standard in practical data processing applications. Both synthetic and field data results confirm that the rational interpolation method significantly outperforms the nonhyperbolic moveout equation method of Alkhalifah and Tsvankin, in combining unbiased parameter (η) estimation with accurate moveout correction, for arbitrary levels of anellipticity and large ODR. The resulting traveltime errors are several orders of magnitude more accurate than the errors obtained from the nonhyperbolic moveout equation of Alkhalifah and Tsvankin, and the method can be tuned to be accurate at any offset range of interest by increasing the order of the interpolation. The improved accuracy at larger ODR should lead to more accurate imaging of steeper dips if the method is used in the context of post or pre-stack time-migration in VTI media.

In the special case of $\eta \lesssim 0.1$ and maximum ODR up to four (we did not test ODR larger than four), the shortcomings in moveout correction and estimated values of η, obtained using the nonhyperbolic moveout equation from Alkhalifah and Tsvankin, are small and probably negligible from a practical point of view. Comparison of the traveltimes used for the rational interpolation (calculated using the above-mentioned customary assumptions) and the shifted hyperbola approximation from Fomel (2004) for Greenhorn-shale anisotropy, showed that the differences between both methods are less than 1 ms, and hence, from a practical point of view, negligible.

Under the customary assumptions that traveltimes of qP-waves in VTI media depend mainly on η and V_{NMO} (i.e., we can choose $\delta = 0$), and that the influence of V_{S0} on traveltimes of qP-waves in TI media is negligible, we found that for a single horizontal VTI layer, the nonhyperbolic moveout curve for fixed values of η and t_0, but varying values of V_{NMO}, can be calculated by simply stretching the offset axis. This observation allows calculation of the traveltimes needed for the rational interpolation from a small number of traveltimes for a few ODR values, some reference value of t_0 (conveniently $t_0^{ref} = 1$ s), and η values ranging from -0.2 to 1.0 in steps of, say, 0.01. These few hundred traveltimes can be precomputed once, in about a minute on a modern PC and subsequently stored in a table. Therefore, the rational interpolation based method has no computational overhead compared to methods based on approximations explicit in the relevant parameters η and V_{NMO}.

Chapter 6

Nonhyperbolic moveout analysis of qP-waves in layered VTI media using rational interpolation[1]

6.1 Summary

The nonhyperbolic moveout observed in common-midpoint gathers is often ascribed to the presence of seismic anisotropy in the subsurface. Therefore, this moveout is commonly used to get (initial) estimates of the relevant seismic parameters describing such anisotropy. For qP-waves in laterally homogeneous transversely isotropic media with a vertical symmetry axis (VTI), these parameters are the anellipticity parameter η and the zero-dip normal moveout velocity V_{nmo}. The τ-p domain is the natural domain for parameter estimation in horizontally layered media since cusps do not occur in this domain and because the horizontal slowness is preserved upon propagation. However, the need for methods that use the τ-p domain to transform the data or to pick traveltimes in the x-t domain, is a practical disadvantage. To overcome this, we combine a τ-p domain inversion technique with rational interpolation of traveltimes in the x-t domain. This combination results in a highly accurate, yet efficient semblance-based method in the x-t domain, for estimating the interval values of η and V_{nmo} from the moveout of qP-waves reflected once in a horizontally layered VTI medium. The efficiency of the method stems from the fact that rational interpolation needs only few (here five) support points to obtain accurate moveout interpolation. The accuracy of the method is verified with a numerical example.

6.2 Introduction

In the past two decades, seismic data processing has gradually developed to allow an estimate of seismic anisotropy to be obtained from the data. A common method to get (initial) estimates of seismic anisotropy is to analyze the nonhyperbolic part of the moveout in a CMP gather (Tsvankin & Thomsen, 1994; Alkhalifah & Tsvankin, 1995; Alkhalifah, 1997; Grechka & Tsvankin, 1998; Tsvankin, 2001, Chapter 4). This analysis originated with the work of Hake *et al.* (1984), who presented a three-term Taylor expansion to describe the

[1]This Chapter is part of a paper in progress intended for submission to *Geophysics* as: Huub Douma and Mirko van der Baan, 2006, Nonhyperbolic moveout analysis of qP-waves in layered VTI media using rational interpolation. An abstract on the work presented in this Chapter has been accepted for presentation at the Annual Meeting of the EAGE as: Huub Douma and Mirko van der Baan, 2006, Semblance-based anisotropy parameter-estimation in layered VTI media using rational interpolation.

moveout in a CMP gather acquired over a horizontally-layered transversely isotropic (TI) medium. Tsvankin & Thomsen (1994) presented an improved rational approximation that provided better accuracy at long offsets, but many authors subsequently pointed out the lack of accuracy at intermediate offsets [e.g., Grechka & Tsvankin (1998), Zhang & Uren (2001), Van der Baan & Kendall (2002), Stovas & Ursin (2004), and Douma & Calvert (2006)]. This observation led to several other methods, such as the $\tau - p$ based method (Van der Baan & Kendall, 2002; Van der Baan, 2004), the shifted-hyperbola-based approach [e.g., Fomel (2004)], continued-fraction-based approaches [e.g., Ursin & Stovas (2006)], and a rational-interpolation-based method (Douma & Calvert, 2006).

Because horizontal slowness is preserved upon propagation through horizontally layered media, and because traveltimes as a function of the horizontal slowness p are single-valued, the τ-p transform is the natural domain for anisotropy parameter estimation in layered media (Hake, 1986; Van der Baan & Kendall, 2002). Since the τ-p transform is a plane-wave decomposition, the relevant velocity is phase velocity rather than group velocity. The latter velocity is mathematically more complex and often follows from approximations to already approximated phase velocities. Hence, anisotropy-parameter estimation in the τ-p domain renders the extra approximation for the group velocity unnecessary, thus leading to more accurate estimates of the relevant parameters.

The disadvantage of the τ-p domain inversion technique is that we either need to (1) transform the x-t gathers to the τ-p domain and pick the τ-p curves in this domain, or (2) pick the $t(x)$ moveout curves using a least-squares fitting procedure and transform these to the τ-p domain. Most interpreters prefer the second option since they have more experience viewing data in the x-t domain (Van der Baan & Kendall, 2002). Manual picking of the moveout curves, however, is in practice cumbersome despite being prone to error.

The approximation to nonhyperbolic moveout of qP waves given by Tsvankin & Thomsen (1994) is a rational approximation that is exact at zero and infinite offset. Recognizing the accuracy of rational approximants together with the need for better accuracy at intermediate offsets, Ursin & Stovas (2006) recently derived a rational approximation (written as a continued fraction) for nonhyperbolic moveout in layered VTI media by matching a Taylor series expansion for the squared traveltime up to sixth order in offset (actually third order in squared offset); such approximation is also known as *Padé approximation* [e.g., Baker (1975) or Bender & Orszag (1978)]. Although this rational approximation provides improved accuracy at intermediate offsets, beyond certain offsets it loses accuracy. To overcome this, but at the same time recognizing the accuracy of the rational approximation, Douma & Calvert (2006) (see Chapter 5) proposed a *rational interpolation* approach to describe nonhyperbolic moveout. In this method, the moveout is obtained through rational interpolation of traveltimes at offsets where accuracy is desired, i.e., at offsets acquired in the field. They showed that for a single horizontal VTI layer, the rational-interpolation-based method, using quadratic polynomials in both the numerator and the denominator, allows for highly accurate moveout correction up to offset-to-depth ratios of about 8, combined with unbiased estimation of V_{hor} (or η).

We extend the rational-interpolation-based method of Douma & Calvert (2006) to a horizontally layered TI medium with a vertical symmetry axis, based on nearly exact

traveltimes derived from the τ-p curve in combination with the acoustic approximation of Alkhalifah (1998). This allows us to set up a three-parameter, i.e., t_0, η, and V_{nmo}, semblance-based analysis in the $t-x$ domain, to invert qP-wave moveout for the interval values of η and V_{nmo} as a function of the zero-offset two-way traveltime t_0, without the use of a Dix-type averaging procedure (Dix, 1955). The presented method is highly efficient due to a combination of expressions for the traveltimes and offsets that are explicit in the interval values of η and V_{nmo} and the horizontal slowness, and because rational interpolation needs only few support points (here five) to obtain accurate moveout interpolation. Since we are dealing with qP-waves only, interpolation can be done in the $x-t$ domain since no cusps can occur on qP wavefronts due to the convexity of the slowness surface of such waves. The method can be extended to pure mode qSV reflections based on the weak anisotropy approximation of the phase velocity for such waves (Thomsen, 1986), provided the measured reflected wavefront does not contain any cusps. In case the reflected wavefront does contain cusps, the $\tau-p$ domain is the natural domain of choice for horizontally layered media, as pointed out by Van der Baan & Kendall (2002).

The organization of this paper is as follows. First, we review the $\tau-p$ curve and subsequently use it to develop the expressions for traveltime and offset of pure-mode qP reflected waves in layered VTI media, as a function of the horizontal slowness. We then proceed to present the rational-interpolation-based procedure for traveltime estimation, and explain its use in interval parameter estimation of η (or V_{hor}) and V_{nmo}. Finally, a numerical example illustrates the method.

6.3 The general $\tau-p$ curve

From simple geometric considerations, it follows that in an (anisotropic) homogeneous medium the horizontal distance $\Delta x'$ traveled by a plane wave within a time interval $\Delta t'$ is given by

$$\Delta x' = v_g^x \Delta t' \,, \tag{6.1}$$

where v_g^x is the horizontal group velocity. The length of the projection of the group-velocity vector on the normal to the wavefront equals the phase velocity [e.g. Vlaar (1968); Helbig (1994, p.23)]. Therefore, in two dimensions,

$$p v_g^x + q v_g^z = 1 \,, \tag{6.2}$$

where p and q are the horizontal and vertical components of the slowness vector $\boldsymbol{p}$, respectively, and v_g^z is the vertical component of the group velocity. Using equation (6.2) in equation (6.1) and solving for $\Delta t'$ we have

$$\Delta t' = p \Delta x' + q \Delta z \,, \tag{6.3}$$

where we used $\Delta z = v_g^z \Delta t'$. The two-way traveltime Δt can be split into two one-way traveltimes $\Delta \acute{t}'$ and $\Delta \grave{t}'$ for the up- and down-going ray, respectively, i.e.,

$$\Delta \acute{t}' = \acute{p}\Delta x' - \acute{q}\Delta z, \quad \Delta \grave{t}' = \grave{p}\Delta x' + \grave{q}\Delta z \,, \tag{6.4}$$

where we denote explicitly the difference in the vertical direction for the up- and down-going rays. For a reflected wave in a horizontal layer, it follows, using Snell's law $\acute{p} = \grave{p} = p$, that

$$\Delta t = \Delta \acute{t}' + \Delta \grave{t}' = p\Delta x + (\grave{q} - \acute{q})\Delta z \,, \tag{6.5}$$

where $\Delta x = 2\Delta x'$ denotes the offset. Assuming that the medium has a horizontal symmetry plane (such as transversely isotropic media with either a vertical or a horizontal symmetry axis, or orthorhombic media with a horizontal symmetry plane, or even monoclinic media with a horizontal symmetry plane) and assuming pure-modes only (i.e., qP-waves, qSV-waves and qSH-waves), we have $\grave{q} = -\acute{q} = q$. Using this in equation (6.5) then gives

$$\Delta t = p\Delta x + 2q\Delta z \,. \tag{6.6}$$

Defining the intercept time $\Delta\tau := \Delta t - p\Delta x$, and replacing the single layer with a stack of homogeneous anisotropic horizontal layers, we obtain the $\tau - p$ curve [e.g., Chapman (2004, section 2.3)]

$$t = px + \tau \,, \tag{6.7}$$

with x denoting the offset and with

$$\tau = \sum_i \Delta t_0^i v_0^i q^i \,, \tag{6.8}$$

where t_0^i is the two-way vertical traveltime in layer i, and v_0^i and q^i are the vertical group velocity and vertical slowness in layer i, respectively. It follows from the definition of the $\tau - p$ curve that

$$\frac{d\tau}{dp} = -x \,. \tag{6.9}$$

6.4 Traveltimes for qP-waves in layered VTI media based on the $\tau - p$ curve

Although equation (6.7) is valid in generally anisotropic horizontally layered media with a horizontal symmetry plane, and valid for qP-waves as well as for qSV-waves and qSH-waves, we treat qP-waves in VTI media only. Since for such waves no cusps can occur in the moveout due to the convexity of the slowness surface, this allows the interpolation to be done in the $x - t$ domain. Also, in horizontally layered VTI media, each vertical plane is a symmetry plane, so the vertical group-velocity is the same as the vertical phase velocity. Therefore, the vertical group velocity v_0^i in equation (6.8) is replaced with the vertical phase velocity V_0^i (i.e., V_{P0}^i for qP-waves).

The traveltimes of qP-waves in laterally homogeneous VTI media depend mainly on

the zero-dip normal-moveout velocity V_{nmo} and the anellipticity parameter η (Alkhalifah & Tsvankin, 1995). The vertical shear-wave phase velocity V_{S0} has negligible influence on the traveltimes of such waves in TI media (Tsvankin & Thomsen, 1994; Tsvankin, 1996; Alkhalifah, 1998). Alkhalifah (1998) derived an approximate relation for the vertical slowness q in terms of V_{nmo} and η by setting $V_{S0} = 0$. This relation is given by

$$V_{P0}^2 q^2 = 1 - \frac{1 - p^2 V_{nmo}^2}{1 - 2\eta p^2 V_{nmo}^2} \,. \tag{6.10}$$

Setting $V_{S0} = 0$ is described as the acoustic approximation by Alkhalifah (1998). Grechka & Tsvankin (1998) showed that the horizontal velocity $V_{hor} = V_{nmo}\sqrt{1+2\eta}$ is better constrained from semblance-based moveout analysis than is η. Therefore, in anticipation of a semblance-based parameter estimation, we rewrite equation (6.10) in terms of V_{hor} and V_{nmo}. When we do this, equation (6.10) becomes

$$V_{P0}^2 q^2 = \frac{1 - p^2 V_{hor}^2}{1 - p^2 \left(V_{hor}^2 - V_{nmo}^2\right)} \,. \tag{6.11}$$

Using equation (6.11) in (6.8) then gives τ explicit in the relevant parameters V_{hor} and V_{nmo}, i.e.,

$$\tau = \sum_i \Delta t_0^i \sqrt{\frac{1 - p^2 \left(V_{hor}^i\right)^2}{1 - p^2 \left[\left(V_{hor}^i\right)^2 - \left(V_{nmo}^i\right)^2\right]}} \,. \tag{6.12}$$

With this expression we can calculate the derivative $d\tau/dp$ and use the resulting expression in equation (6.9) to find the offset x explicit in V_{hor} and V_{nmo}. Doing this gives

$$x = \sum_i \Delta t_0^i \left(\frac{p \left(V_{nmo}^i\right)^2 / \sqrt{1 - p^2 \left(V_{hor}^i\right)^2}}{\left\{1 - p^2 \left[\left(V_{hor}^i\right)^2 - \left(V_{nmo}^i\right)^2\right]\right\}^{3/2}} \right) \,. \tag{6.13}$$

Finally, using equations (6.12) and (6.13) in (6.7), we get an expression for the traveltime as function of the horizontal slowness p explicit in the relevant parameters V_{hor}^i and V_{nmo}^i, i.e.,

$$t = \sum_i \Delta t_0^i \left(\frac{p^2 \left(V_{nmo}^i\right)^2 / \left\{1 - p^2 \left[\left(V_{hor}^i\right)^2 - \left(V_{nmo}^i\right)^2\right]\right\} + \left[1 - p^2 \left(V_{hor}^i\right)^2\right]}{\sqrt{\left[1 - p^2 \left(V_{hor}^i\right)^2\right] \left\{1 - p^2 \left[\left(V_{hor}^i\right)^2 - \left(V_{nmo}^i\right)^2\right]\right\}}} \right) \,. \tag{6.14}$$

The weak anisotropy approximation (Thomsen, 1986) can be used to linearize the phase velocity V_{SV} for qSV waves as a function of phase angle, which can subsequently be written [equation (31) from Van der Baan & Kendall (2002)] in terms of the horizontal

slowness as

$$V_{SV}(p) \approx \frac{-1 + 2\sigma p^2 V_{S0}^2 + \left\{ \left(1 - 2\sigma p^2 V_{S0}^2\right)^2 + 8\sigma p^4 V_{S0}^4 \right\}^{1/2}}{4\sigma p^4 V_{S0}^2}, \tag{6.15}$$

with $\sigma := (V_{P0}/V_{S0})^2(\epsilon - \delta)$. Then the vertical slowness q can be found using $q^2 = 1/V_{SV}^2 - p^2$, which can subsequently be used in equation (6.8) to find the intercept time. Substitution of this result in equations (6.7) and (6.9) gives the equivalent expressions to equations (6.13) and (6.14) for the offsets and associated traveltimes of reflected qSV waves. The resulting expressions could be made explicit in the NMO velocity V_{nmo}^{SV} using the relation $V_{nmo}^{SV} = V_{S0}\sqrt{1 + 2\sigma}$. Such expressions could be used only in an $x - t$ based interpolation scheme provided the wavefront does not contain any cusps. Although such a treatment is possible, linearized approximations for qSV-waves are often not very accurate because the presence of the squared V_{P0}/V_{S0} in the definition of σ often makes σ much larger than ϵ and δ (Tsvankin, 2001, p.26-27). Therefore we refrain from this treatment and subsequent numerical verification.

6.5 Rational interpolation

Equations (6.13) and (6.14) are explicit in p. For an x-t based semblance analysis, however, we require traveltimes for all offsets acquired in the field. Finding the p value associated with a certain offset x in equation (6.13) corresponds to a simplified 2-point ray-tracing problem, which can be solved using a straightforward bisection approach since the offset x and traveltime t in equations (6.13) and (6.14) are single-valued and monotonically increasing with increasing p-value. In principle this can be done for each available offset and the resulting moveout curves can be used to perform a semblance-based parameter estimation in the $x - t$ domain. Here, however, we propose a more elegant and computationally less demanding procedure that uses the traveltimes and offsets calculated for a few p-values only in the rational-interpolation-based approach of Douma & Calvert (2006) (see Chapter 5). In this way, the method of Douma & Calvert, which is based on a single horizontal VTI layer, can be extended to horizontally layered VTI media. The proposed method naturally encompasses the single-layer case, hence removing the need for the tabulated approach used by Douma & Calvert.

A rational approximation to a function $t(x)$ is generally written as [e.g., Stoer & Bulirsch (1993, p.58-63)]

$$t(x) \approx \frac{N_L(x)}{D_M(x)} = \frac{n_0 + n_1 x + \cdots + n_L x^L}{d_0 + d_1 x + \cdots + d_M x^M}, \tag{6.16}$$

with $N_L(x)$ a polynomial of order L, and $D_M(x)$ a polynomial of order M. We denote such an approximation as $[L/M]$. This rational approximation is determined by the $L + M + 2$ unknown coefficients n_i, $i = 0, 1, \ldots, L$ and d_i, $i = 0, 1, \ldots, M$. Since these coefficients can be determined up to only a common factor $\rho \neq 0$, the $[L/M]$ rational approximation is fully

determined by $L+M+1$ support points (x_i, t_i) satisfying

$$t(x_i) = \frac{N_L(x_i)}{D_M(x_i)} = t_i \; . \tag{6.17}$$

Hence, it is necessary that the coefficients satisfy

$$N_L(x_i) - t_i D_M(x_i) = 0 \; . \tag{6.18}$$

Satisfying equation (6.18), however, is not a sufficient condition for equation (6.17) to hold since $N_L(x_i)$ and $D_M(x_i)$ may both be zero. In that case, the rational interpolation problem is unsolvable, and support points (x_i, t_i) cannot be reached and are thus missed by the rational function $N_L(x)/D_M(x)$. Such support points are usually referred to as *inaccessible*. If x_i is an inaccessible point, the functions $N_L(x)$ and $D_M(x)$ have (at least) a common factor $x - x_i$. If $N_L(x)$ and $D_M(x)$ have no such common factor, the rational interpolation problem is solvable (Stoer & Bulirsch, 1993, p.58-63).

If the $L+M+1$ support points (x_i, t_i) of an $[L/M]$ rational approximant can be interpolated by a $[L'/M']$ rational approximation with $L'+M' < L+M$, the support points are said to be in *special position* (Stoer & Bulirsch, 1993, p.61-62). Suppose $i_1, \ldots, i_\alpha$ are the subscripts of the inaccessible points of the $[L/M]$ rational approximant $t(x)$ in equation (6.16). Then, $N_L(x)$ and $D_M(x)$ have common factor $\Pi_{j\in\{i_1,\ldots,i_\alpha\}}(x - x_j)$, which after cancellation leaves an equivalent $[L'/M']$ rational approximant with $L'+M'+1 = L+M+1-2\alpha$. Since the remaining polynomials $N'_L(x)$ and $D'_M(x)$ have no more common factors, they form a rational approximation $t'(x) = N'_L(x)/D'_M(x)$ that solves the rational interpolation problem through the remaining $L+M+1-\alpha$ accessible points. Since $L+M+1-\alpha > L+M+1-2\alpha = L'+M'+1$, it follows that the accessible support points of a nonsolvable rational interpolation problem can always be interpolated by a $[L'/M']$ rational approximation with $L'+M' < L+M$, and are thus always in special position. This means that the nonsolvability of a rational interpolation problem is really a matter of degeneracy. Therefore, the possible nonsolvability of a rational interpolation problem is never an issue in practice, since it can be overcome by adding arbitrarily small perturbations to the support points.

If x and t in equation (6.16) denote squared offset and squared traveltime, respectively, it follows that for purely hyperbolic moveout all the support points are in special position. As shown in Appendix N, this means that in that case the linear system (6.18) is degenerate. This was also noted by Douma & Calvert (2006), who observed that this degeneracy can be overcome by arbitrarily small perturbations (in the form of numerical noise) of the support points (x_j, t_j). If, however, x and t denote offset and time instead of squared offset and squared time, then there is no degeneracy for hyperbolic moveout. Here, we therefore interpolate the traveltimes as a function of offset as opposed to the squared traveltimes as a function of the squared offsets, as done by Douma & Calvert (2006).

Douma & Calvert explicitly solve the linear system (6.18) for [2/2] rational interpolation in their appendix A (see Appendix M of this thesis). Alternatively, the rational approximant $N_L(x)/D_M(x)$ can be written as a Thiele continued fraction (Stoer & Bulirsch,

1993, p.63-67) given by

$$\frac{N_L(x)}{D_M(x)} = t_0 + \frac{x-x_0}{\rho(x_0,x_1)+} \frac{x-x_1}{\rho(x_0,x_1,x_2)-\rho(x_0)+} \cdots \frac{x-x_{M+L-1}}{\rho(x_0,\ldots,x_{M+L})-\rho(x_0,\ldots,x_{M+L-2})}, \tag{6.19}$$

where we use the notation $\frac{a}{b+}\frac{c}{d} = \frac{a}{b+c/d}$. Here, the reciprocal differences ρ are defined by

$$\rho(x_i) := t_i, \tag{6.20}$$

$$\rho(x_i,x_k) := \frac{x_i-x_k}{t_i-t_k}, \tag{6.21}$$

$$\rho(x_i,x_{i+1},\ldots,x_{i+k}) := \frac{x_i-x_{i+k}}{\rho(x_i,\ldots,x_{i+k-1})-\rho(x_{i+1},\ldots,x_{i+k})} + \rho(x_{i+1},\ldots,x_{i+k-1}). \tag{6.22}$$

Here, we use the same order of rational interpolation that Douma & Calvert used, i.e., a [2/2] rational interpolation, since they showed that this order of interpolation (at least for a single horizontal homogeneous VTI layer) provides high accuracy in the traveltimes up until offset-to-depth ratios of 8. Setting $L = M = 2$ in equation (6.19) gives the resulting Thiele continued-fraction interpolation, i.e.,

$$t(x) \approx \frac{N_2(x)}{D_2(x)} = t_0 + \frac{x-x_0}{\rho(x_0,x_1)+} \frac{x-x_1}{\rho(x_0,\ldots,x_2)-\rho(x_0)+} \frac{x-x_2}{\rho(x_0,\ldots,x_3)-\rho(x_0,x_1)+} \frac{x-x_3}{\rho(x_0,\ldots,x_4)-\rho(x_0,\ldots,x_2)}. \tag{6.23}$$

This equation can be used to find the function values $t(x)$ at values of x different from the support abscissas x_i. Note that for the [2/2] rational interpolation *only four* support points need to be calculated, because t_0 is treated as a parameter, rendering the method highly efficient.

6.6 Three-parameter semblance-based parameter estimation in the $x-t$ domain

We use the expressions for the traveltimes and associated offsets, i.e., equations (6.14) and (6.13), respectively, to calculate the support points for the rational interpolation. That is, given four p-values p_j ($j = 1,\ldots,4$), and given the model parameters Δt_0^i, V_{hor}^i, and V_{nmo}^i, we calculate the pairs (x_j,t_j). Together with the pair $(x_0, t_0 = \sum_i \Delta t_0^i)$, these pairs are the necessary five support points for the [2/2] rational interpolation that can be used to calculate the reciprocal differences from equations (6.20)-(6.22). Using the resulting reciprocal

differences, we then have the necessary ingredients to calculate the Thiele continued-fraction interpolation using equation (6.23). The moveout curve obtained in this fashion, can be used either to perform a semblance analysis or to moveout correct a CMP gather.

For a single horizontal VTI layer, Douma & Calvert (2006) used 5 regularly spaced offsets to act (together with the associated traveltimes) as the support points of the [2/2] rational interpolation. For the layered case, we obtain the four[2] necessary p-values using a bisection approach in combination with equation (6.13). That is, given four offsets (or offset-to-depth ratios), we find four horizontal slownesses p_j that have offsets close to the prescribed ones (say within an accuracy of 100 m) using bisection. For a single horizontal layer this bisection approach is in principle unnecessary, since the horizontal slowness as a function of offset [derived from equation (6.13)] can be shown to satisfy a quartic equation that can be solved explicitly (see Appendix G for the general solution to a quartic equation). Alternatively, we could simply use four equally spaced p_j values, with the largest p_j equal $1/\left(\max_i V^i_{hor}\right)$. This is computationally the most efficient option. We have not tested this approach, however, since we aim to stay close to the method of Douma & Calvert (2006) which obtains accurate interpolation using regularly spaced offset-to-depth ratios. Moreover, because the rational interpolation achieves high accuracy with so few support points, the computational overhead of bisection is negligible.

Since the support points are explicit in V^i_{hor}, V^i_{nmo}, and Δt^i_0, a semblance-based inversion in the x-t domain is feasible as a function of the three parameters t_0, V^i_{nmo}, and V^i_{hor}. In this way the best-fit values of V^i_{hor} and V^i_{nmo} can be determined as a function of t_0, by finding the maximum semblance combinations of V^i_{hor} and V^i_{nmo} as a function of t_0. For pure-mode qSV waves in the absence of cusps, a similar semblance analysis is feasible based on the three relevant parameters t_0, σ_i, and $V^i_{nmo,SV}$, where σ_i denotes the interval σ.

In principle, a semblance-based global inversion can be set up to determine the values of V^i_{hor} and V^i_{nmo} for all layers at the same time, or by applying the inversion in a layer-stripping fashion. In practice, however, both approaches are often not desirable, because of the trade-off between V^i_{hor} and V^i_{nmo} for limited offset acquisition geometries [e.g, Alkhalifah (1997); Grechka & Tsvankin (1998). This tradeoff makes the inversion sensitive to errors in the observed traveltimes caused by the presence of noise or imperfect static corrections. Because this tradeoff is reduced with increasing offset-to-depth ratio (Alkhalifah, 1997), however, such procedures could be used when larger offset-to-depth ratios (approximately four or larger) are available in the data. When such data are not available and if the layer of interest is known, the overburden can be modeled as a single layer with certain effective values of V_{hor} and V_{nmo} [e.g., Tsvankin (2001, section 4.1.3)]. Once these values are established, the values of V^i_{hor} and V^i_{nmo} can be estimated by finding the maximum semblance value over a range of values of V^i_{hor} and V^i_{nmo}, while keeping the effective values of V_{hor} and V_{nmo} of the overburden fixed. This method would reduce the problem of error-propagation in a layer-stripping approach when large offset-to-depth ratios (say ≥ 3) are not available in the data.

[2]Because the pair (x_0, t_0) simply relates to $p_0 = 0$, we need only calculate four p_j values, instead of five.

Table 6.1. Model parameters used in ray-tracing to determine the traveltimes from a layered model. The associated CMP gather is shown in Figure 6.1a.

layer	h (m)	V_{nmo} (m/s)	V_{hor} (m/s)	η	V_{P0} (m/s)	V_{S0} (m/s)	ϵ	δ
1	1000	2098	2098	0.00	2000	300	0.050	0.05
2	2000	2000	2298	0.16	2000	300	0.160	0.00
3	3000	2892	3747	0.34	3048	300	0.255	-0.05
4	4000	2464	3882	0.74	3292	300	0.195	-0.22

6.7 Numerical example

Douma & Calvert (2006) (see Chapter 5) applied the [2/2] rational interpolation method for a single horizontal VTI layer to a layered model to obtain effective value of η and V_{nmo}. Here, we test our method for interval parameter estimation of η and V_{nmo} using the same layered model that Douma & Calvert used. The model parameters are given in Table 6.1. Figure 6.1a shows a CMP gather for this model. The top layer is elliptically anisotropic, i.e., $\eta = 0$, whereas the remaining layers have subsequently increasing values of η. The values of V_{nmo}, V_{hor}, and η for each layer are listed in Figure 6.1a. The level of anellipticity for the third layer (i.e., $\eta = 0.34$) can be considered high, whereas the anellipticity in the fourth layer (i.e., $\eta = 0.74$) is extreme. The values of V_{nmo}, V_{hor}, and η, for the second layer, correspond to the model parameters used in Figures 1 and 2 of Grechka & Tsvankin (1998)[3]. For the third and fourth layer, the model parameters correspond to a shale under zero confining pressure, and to Green River shale, respectively, [see Table 1 in Thomsen (1986) for these two cases]. Hence, even though the anellipticity in the third and fourth layer are high, they have been observed in laboratory measurements.

The CMP gather shown in Figure 6.1a was obtained using anisotropic ray-tracing and subsequently setting all amplitudes to 1.0. Hence, realistic amplitude and phase variations with offset are not modeled in this gather. It is known that semblance based methods can to some extent suffer from amplitude variations with offset [e.g., Sarkar *et al.* (2002)]. Here, we eliminate this complication in order to focus our attention on the ability of the proposed method to accurately determine the interval parameters of V_{nmo}, V_{hor}, and η from the traveltimes in CMP gathers.

We want to highlight that the proposed method is accurate up to large offset-to-depth ratios. Therefore, for all events in Figure 6.1a, the maximum offset-to-depth ratio is four. Knowing that for such large ratios the tradeoff between V_{nmo} and V_{hor} is small, we employ

[3]These parameters were originally chosen because such values of η were observed on field data (Alkhalifah *et al.*, 1996).

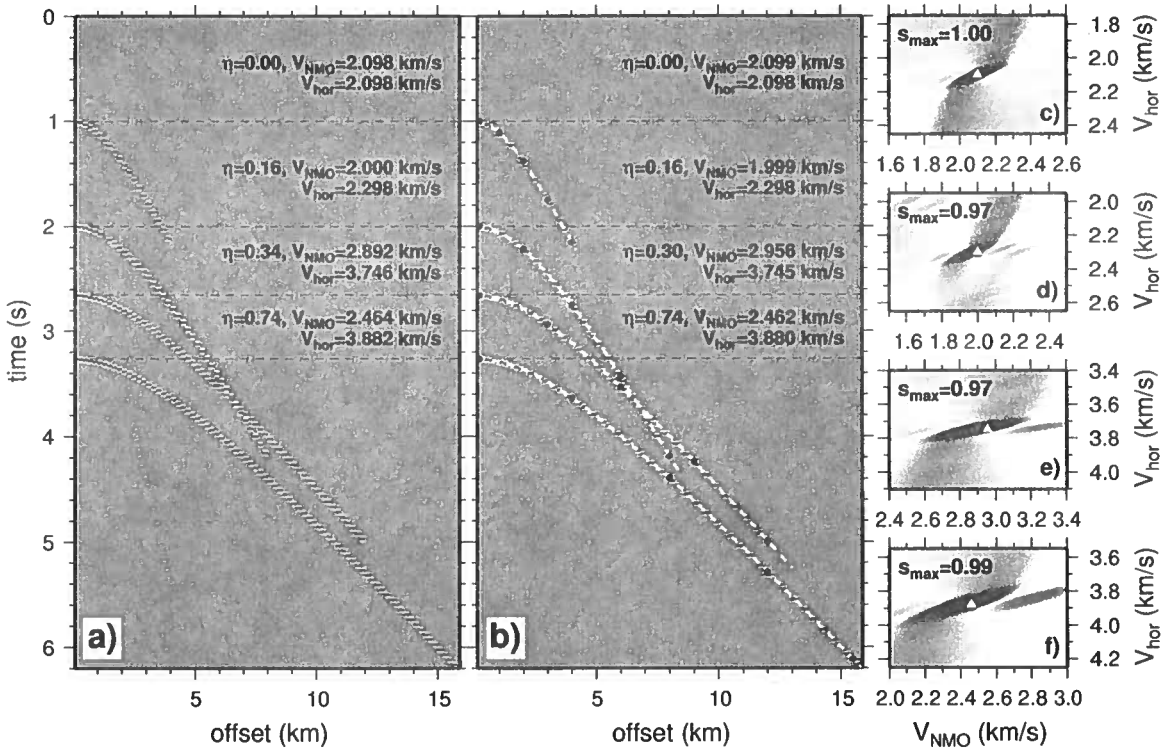


Figure 6.1. (a) CMP gather with qP-reflections from a 4-layer model generated using ray-tracing (see Table 6.1 for parameters). The layers are outlined by dashed lines, and the true values of η, V_{hor} and V_{nmo} are indicated. (b) As (a) but with estimated values of η, V_{hor} and V_{nmo} indicated together with their associated traveltimes (dashed white lines). The support points are marked by the black dots. (c-f) Semblance scans as a function of the interval values of V_{hor} and V_{nmo} for layer 1 (c) through 4 (f), respectively. The grey circles indicate the true values of associated interval values of V_{hor} and V_{nmo} while the white triangles indicate the interval values obtained from the maximum semblance ($s_{\max}$).

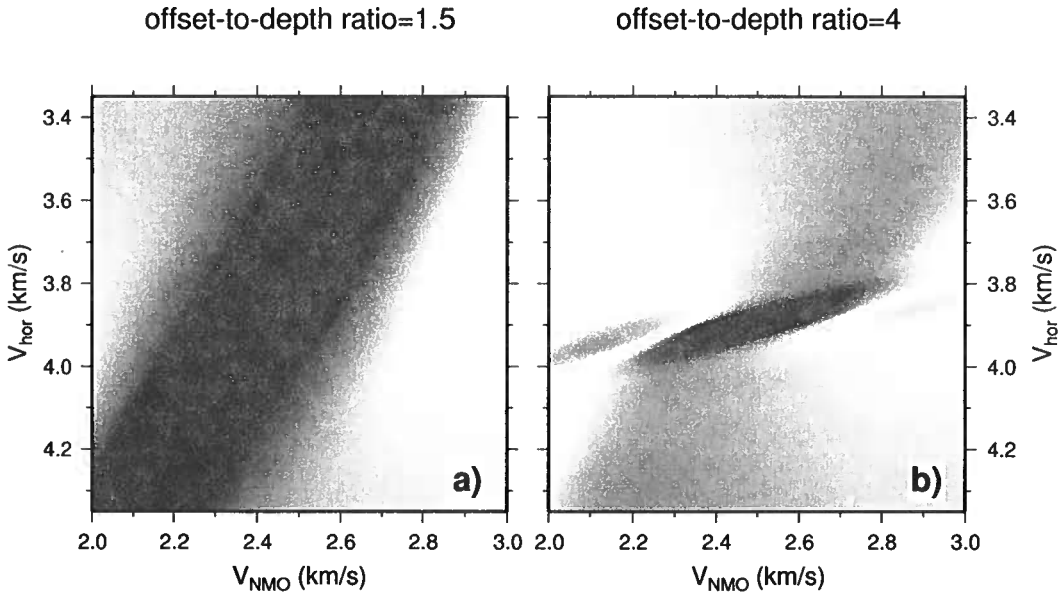


Figure 6.2. Semblance scans as a function of the interval values of V_{hor} and V_{nmo} for layer 4 for an offset-to-depth ratio of 1.5 (a) and 4 (b).

the proposed method in a layer-stripping setting. Figures 6.1c-f show the semblance scans from the first through the fourth event in Figure 6.1a as a function of the *interval* values V^i_{nmo} and V^i_{hor}. For each layer, the semblance scans were obtained by fixing the overburden values of V^i_{nmo} and V^i_{hor} to their estimated values. Then, by varying the interval values V^i_{nmo} and V^i_{hor} in the layer of interest and calculating the traveltime curves for each combination of V^i_{nmo} and V^i_{hor} using the rational-interpolation method given above, we determined the semblance values along these curves. The location of the maximum semblance value is indicated by the white triangle, while the location corresponding to the true model values of V^i_{nmo} and V^i_{hor} is indicated by the grey circle. Note the excellent agreement between the obtained values of V^i_{nmo} and V^i_{hor} (and thus η_i) and the true interval values of these parameters (see Figure 6.1b), as inferred from the close proximity of the locations of the white triangle and the grey circle. Figure 6.1b shows the traveltime curves (calculated using the rational-interpolation approach) associated with the obtained interval values V^i_{nmo} and V^i_{hor} (white dashed lines) superimposed on the events in the CMP gather, confirming the close match between the true traveltimes and the traveltimes from the inverted interval parameters V^i_{nmo} and V^i_{hor}. This close match is confirmed by the high maximum semblance values s_{max} (shown in Figures 6.1c-f), indicating the excellent agreement between the actual traveltimes and the obtained traveltimes, even for the large offset-to-depth ratio of four. The support points used for the rational interpolation are indicated by the black circles.

From the shapes of the semblance scans for all layers, it follows that V^i_{hor} is better resolved than V^i_{nmo}. This can be attributed to the relatively large value of four used for the maximum offset-to-depth ratio. Indeed, larger offsets put tighter constraints on the

Table 6.2. Effective values of V_{nmo}, V_{hor}, and η for the layered model shown in Figure 6.1a obtained using the Alkhalifah-Tsvankin approximation. The maximum offset-to-depth ratio used to obtain these values is 1.5.

layer	V_{nmo} (m/s)	V_{hor} (m/s)	η
1	2096	2100	0.00
2	2047	2193	0.07
3	2284	2816	0.26
4	2328	2995	0.33

inversion of V_{hor} than do the smaller offsets [e.g., Alkhalifah (1997)]. Figure 6.2 confirms this in the context of the interval value V^i_{hor}, by showing the semblance scan for the fourth layer for a maximum offset-to-depth ratio of 1.5 (Figure 6.2a) and 4 (Figure 6.2b). Note that the larger offsets do not really add any constraints on V^i_{nmo}, since V^i_{nmo} is mainly determined by the near offsets. Therefore, the small discrepancies between the true values and obtained interval values η_i are mainly caused by small deviations of the obtained values of V_{nmo} from their true values (see Figure 6.1). We attribute these differences to the use of the acoustic approximation to determine the support points.

Grechka & Tsvankin (1998) introduced a method to find the interval values V^i_{nmo} and η_i from the *effective* values of V_{nmo} and V_{hor} obtained from the Tsvankin-Thomsen approximation (Tsvankin & Thomsen, 1994) [later rewritten in terms of η by Alkhalifah & Tsvankin (1995)]. This approximation describes the nonhyperbolic moveout of qP-waves in a single horizontal VTI layer. The effective values of V_{nmo} and V_{hor} are obtained by finding the best-fit values of V_{nmo} and V_{hor} using this approximation on a CMP gather containing data from a (horizontally) layered VTI medium. Because this approximation lacks accuracy at larger offset-to-depth ratios and as a result introduces a bias in the estimated values of η (see Chapter 5), we compare the proposed method to that of Grechka & Tsvankin (1998, equations (14), (16) and (17)) for a maximum offset-to-depth ratio of 1.5 only. Table 6.3 summarizes the results, while Table 6.2 shows the effective values of V_{nmo}, V_{hor} (and η) used to calculate the interval estimates shown in Table 6.3; the effective values were obtained through a (maximum) semblance analysis based on the Thomsen-Tsvankin approximation (Tsvankin & Thomsen, 1994) for a single horizontal VTI layer. In Table 6.3 the heading "GT" denotes the method by Grechka and Tsvankin, while the heading RI denotes the proposed method based on rational interpolation. The proposed method achieves higher accuracy in the interval values V^i_{nmo}, V^i_{hor}, and η^i than does the method of Grechka and Tsvankin, even for high to extreme levels of anellipticity. The small deviations of V^i_{hor} (and thus η^i), obtained with the rational-interpolation method, from their true values, are due mainly to the lack of resolution in the horizontal velocity of the data for small offset-to-

depth ratios in combination with the use of the acoustic approximation in deriving equations (6.13) and (6.14).

Table 6.3. Comparison of actual interval values of V_{nmo}, V_{hor}, and η, and the values obtained using the inversion formulas of Grechka & Tsvankin (1998) and the presented method based on rational interpolation, for a maximum offset-to-depth ratio of 1.5.

	V_{nmo} (m/s)			V_{hor} (m/s)			η		
layer	Actual	GT	RI	Actual	GT	RI	Actual	GT	RI
1	2098	2096	2098	2098	2100	2096	0.00	0.00	0.00
2	2000	1997	2003	2300	2548	2290	0.16	0.31	0.15
3	2892	2886	2918	3745	3721	3733	0.34	0.33	0.32
4	2460	2511	2490	3880	3583	3851	0.74	0.52	0.70

6.8 Discussion

Rational approximations have achieved relatively accurate traveltime approximations both here and in previous studies. In light of the reported accuracy and efficiency, it therefore seems reasonable to speculate that the underlying functional dependence of the traveltime curves in layered VTI media is close to rational. This would explain the high accuracy of the proposed method with only few (here five) support points.

Rational interpolation can be accompanied by the presence of (nearby)[4] unwanted poles on (or near) the interval of interpolation [e.g., Berrut & Mittelmann (1997, p.357) or Berrut & Mittelmann (2000)]. Such poles occur when the denominator becomes close to zero away from the support abscissas x_i. After the reciprocal differences in the Thiele continued fraction are calculated, the presence of such poles can easily be verified by checking for zeroes in the denominator of the corresponding [2/2] rational approximation. Also, because in the absence of nearby poles in the complex plane the derivative of the traveltime curve should always be positive in layered VTI media, the presence of such poles can be detected by checking for a change in sign of the derivative of the rational approximation $t(x)$. In case a pole is identified, the support points can be perturbed slightly using noise injection (i.e., small perturbation of the support points), to remove the pole. Alternatively, support

[4]Here, nearby is meant in the complex plane; in practice nearby would likely be defined as having a distance (in the complex plane) smaller than the smallest distance between the support abscissas x_i and x_{i+1}.

points at different offsets (x_i) can be used. In the unlikely event that the poles are persistent under these perturbations, the rational interpolation could be replaced with a cubic spline interpolation, at the cost of the need for more support points to obtain (sub-millisecond) accuracy comparable to that of the [2/2] rational interpolation. In this work we have not observed the presence of such unwanted poles.

The presented method shows that with only few (five) support points, highly accurate traveltime curves as a function of offset can be obtained using [2/2] rational interpolation. Even though we only show its accuracy in the context of moveout analysis of qP-waves in layered VTI media, it seems worthwhile to further investigate the use of rational interpolation in more general traveltime calculations, in particular in the context of sparse storage of traveltimes. Such sparse storage is relevant in everyday seismic data processing where large traveltime tables are being used for imaging and inversion.

6.9 Conclusion

Because the horizontal slowness is preserved upon propagation through horizontally layered media, and because the traveltimes as a function of the horizontal slowness p are single-valued (i.e., no cusps), the τ-p transform is the natural domain for anisotropy parameter estimation in layered media. However, quality control of processing results and interpretation of seismic data is preferably done in the t-x domain. Based on the $\tau - p$ curve and the acoustic approximation of Alkhalifah (1998), we have presented expressions for the traveltimes and associated offsets of qP-waves in horizontally layered VTI media that are explicit in the interval values of V_{hor} and V_{nmo} and the horizontal slowness p. We have shown that using these expressions in combination with an $x - t$ based rational interpolation procedure, leads to highly accurate moveout curves for qP-waves in horizontally layered VTI media. These curves are in turn used to set up a three-parameter semblance analysis in the $x - t$ domain for the estimation of V_{nmo} and V_{hor} (and thus η) as a function of t_0. This method overcomes the need to process the data in the τ-p domain or to pick traveltimes in the x-t domain, as in the method of Van der Baan & Kendall (2002). The presented method extends the rational interpolation based method of Douma & Calvert (2006), which is valid for a single horizontal VTI layer, to horizontally layered VTI media. Since this method naturally encompasses the single-layer case, it removes the need for the tabulated approach used by Douma & Calvert (2006). The efficiency of the method stems from the fact that only few (four) support points need to be calculated to achieve highly accurate moveout curves with rational interpolation.

Acknowledgment

The anisotropic ray-tracing used to calculate the traveltimes for the events in the CMP gather of Figure 6.1a, was done using the ANRAY raytracing package from the Seismic Waves in Complex 3D Structures (SW3D) consortium project at Charles University in Prague (http://sw3d.mff.cuni.cz/).

Chapter 7

Correcting for bias due to noise in coda wave interferometry[1]

7.1 Summary

Coda wave interferometry (CWI) utilizes multiply scattered waves to diagnose small changes in a medium by using the scattering medium as an interferometer. Since the medium is usually stationary over the duration of a seismic experiment, different (non-overlapping) time windows in the coda allow for independent estimates of the medium perturbation. If the seismograms are contaminated with noise, only those time windows can be used for which the amplitude of the coda is significantly above the ambient noise level. This limits the usable number of independent time windows. Here, we show how bias due to noise in CWI can be accounted for, by deriving a correction factor for the cross-correlation coefficient, under the assumptions that the stochastic processes underlying the noise realizations in both signals are mutually uncorrelated and stationary with zero mean. This correction factor allows more time windows further into the decaying coda to be used, and hence allows for a reduction of the error bars on the medium perturbation estimates. We demonstrate the validity of this correction factor by using data from a numerical experiment and field measurements. These experiments involve the displacement of point scatterers and a change in the source location, respectively. The application of our correction factor is not limited to CWI, but can be used to correct for bias induced by noise in any application that uses cross-correlation between different signals that are contaminated with noise.

7.2 Introduction

Multiply scattered wavefields have been experimentally shown to be remarkably stable with respect to perturbations of the boundary conditions of experiments with multiply scattered waves (Derode *et al.*, 1995, 1999). Due to this stability, the information carried by multiply scattered waves has been successfully used in an industrial context (e.g., Fink, 1997). Coda wave interferometry (Roberts *et al.*, 1992; Snieder *et al.*, 2002; Snieder, 2002, 2004; Grêt *et al.*, 2006) uses multiply scattered waves to detect small changes by using the scattering medium as an interferometer. Since multiply scattered waves dominate the final

[1]This Chapter has been published as: Huub Douma and Roel Snieder, 2006, Correcting for bias due to noise in coda wave interferometry, Geophysical Journal International, **164**, pp. 99-108.

portions of a seismogram, they are usually referred to as coda waves just as, in musical notation, the coda denotes the closing part of a musical piece. Hence the name *coda wave interferometry* (CWI). Since CWI uses multiply scattered waves, it is inherently more sensitive to changes in the medium than are techniques based on single scattering, as multiply scattered waves sense changes in the medium multiple times.

In parallel, but independently, diffusing acoustic wave spectroscopy (DAWS) (Page *et al.*, 2000; Cowan *et al.*, 2002) was developed as the classical equivalent of diffusing wave spectroscopy (DWS) (Maret & Wolf, 1987; Pine *et al.*, 1988; Yodh *et al.*, 1990; Weitz & Pine, 1993). In DWS light is used to study different aspects of strongly scattering media, whereas in DAWS classical waves are used to probe such media. DWS has been used in many applications such as, e.g., determining the aging of foams, particle sizing, and determining the motion of particles in fluidized suspensions on ångström length scales (Weitz & Pine, 1993). So far, DAWS has mainly been used to determine the relative mean square displacement of fluidized suspensions of particles (Page *et al.*, 1999; Cowan *et al.*, 2000; Page *et al.*, 2000; Cowan *et al.*, 2002). CWI has been succesfully used to measure the nonlinear dependence of seismic velocity in rocks on temperature (Snieder *et al.*, 2002), to monitor volcanos (Grêt *et al.*, 2005), and to estimate source displacement (Snieder & Vrijlandt, 2005). The ability to use CWI to determine the relative mean square displacement of point scatterers from noise-free waveforms was established using a numerical experiment by Snieder *et al.* (2002). Both CWI and DAWS use the amplitude information as well as the phase information of the multiply scattered wavefields, and are both based on a path summation approach to model the multiply scattered wavefields. Hence, both methods are in principle the same, but have been used for different applications.

CWI is based on a measure of cross-correlation between multiply scattered wavefields recorded before and after a medium has changed. The cross-correlation coefficients calculated for different (non-overlapping) time windows provide independent estimates of the medium perturbations. These independent estimates in turn allow for the calculation of error estimates of the perturbation; the larger the number of independent (unbiased) measurements, the smaller the error estimates.

When the coda is contaminated with noise, the number of independent time windows that can be used is limited to traveltimes where the ambient noise level is small compared to the amplitudes of the multiply scattered waves. To be able to use as many independent time windows as possible, and hence reduce the error bars on the inferred perturbation, it is important to correct the cross-correlation function for the bias caused by noise. In this paper, we show how this bias in CWI can be corrected for, by deriving a correction factor for the cross-correlation coefficient based on the assumptions that the stochastic processes underlying the noise realizations in both signals are mutually uncorrelated and stationary with zero mean. We demonstrate its validity by using data from a numerical experiment involving the displacement of point scatterers, and field data involving the displacement of the source location. The organization of this paper is as follows. We briefly state the essence of CWI and derive the noise correction factor. Subsequently we show the validity of the correction factor using numerical and field experiments. We conclude with a short discussion of the results. In Appendix O we derive a condition for the reliability of the

correction factor for the zero-lag cross-correlation coefficient. This condition can be used to determine the time windows in the coda where the correction factor is reliable.

7.3 Coda wave interferometry

Here, we summarize the essence of CWI, but refer to Snieder (2006) and references therein, for a treatment of the theory leading to the expressions presented below.

In CWI, we measure the normalized cross-correlation between seismograms $u_u(t)$ and $u_p(t)$, where $u_u(t)$ and $u_p(t)$ are the wavefields recorded at the same location from an unperturbed and perturbed medium, respectively. We aim to infer the level of the perturbation in the medium from this correlation. The normalized time-shifted cross-correlation is calculated as the time-windowed correlation coefficient $r(t_s;t,t_w)$, given by

$$r(t_s;t,t_w) := \frac{(u_u,u_p)_{(t_s;t,t_w)}}{\sqrt{(u_u,u_u)_{(t_s=0;t,t_w)}\,(u_p,u_p)_{(t_s=0;t,t_w)}}}\,, \tag{7.1}$$

with

$$(u_u,u_p)_{(t_s;t,t_w)} := \frac{1}{2t_w}\int_{t-t_w}^{t+t_w} u_u(t')u_p(t'+t_s)dt'\,,$$

where t_s is the time-shift in the cross-correlation, t is the central window time, and t_w is half the duration of the time window. If we assume that the perturbation in the medium mainly influences the traveltimes of the waves, with a resulting time shift τ_T for the wave that travels along trajectory T (Snieder, 2006), the maximum value of the correlation coefficient, $\max[r(t_s;t,t_w)]$, is related to the variance of the traveltime perturbations σ_τ^2 by

$$\max[r(t_s;t,t_w)] = 1 - \frac{1}{2}\overline{\omega^2}\sigma_\tau^2\,, \tag{7.2}$$

where the frequency $\overline{\omega^2}$ is given by

$$\overline{\omega^2} := \frac{\int_{t-t_w}^{t+t_w}\dot{u}_u^2(t')dt'}{\int_{t-t_w}^{t+t_w}u_u^2(t')dt'}\,, \tag{7.3}$$

and σ_τ^2 is defined as

$$\sigma_\tau^2 := \frac{\sum_T A_T^2\,(\tau_T - \langle\tau\rangle)^2}{\sum_T A_T^2}\,. \tag{7.4}$$

In equations (7.3) and (7.4), $\dot{u}$ denotes the time derivative of $u(t)$, A_T^2 is the intensity of the wave that has traveled along trajectory T, and $\langle\tau\rangle$ denotes the intensity-weighted traveltime perturbation averaged over all scattering paths. It turns out that for the two types of perturbation studied in this paper, i.e., small displacement of the scatterers and a change in the source location, the assumption that the perturbation mainly influences the traveltimes, is appropriate. Note that in equation (7.2), as well as throughout the remainder of this paper, the maximum is calculated with respect to t_s for given values of t and t_w.

7.4 Correcting for the bias due to noise

In equation (7.2), the waveforms $u_u(t)$ and $u_p(t)$ used to calculate $\max[r(t_s;t,t_w)]$ are assumed to be free of noise. To study the influence of noise on CWI, we derive a correction factor for the correlation coefficient when the waveforms $u_u(t)$ and $u_p(t)$ are contaminated with noise. We define

$$u'_u(t) := u_u(t) + n_u(t)\,, \quad u'_p(t) := u_p(t) + n_p(t)\,, \tag{7.5}$$

where $u'_u(t)$ and $u'_p(t)$ are the noise-contaminated signals, $u_u(t)$ and $u_p(t)$ the noise-free waveforms, and $n_u(t)$ and $n_p(t)$ the noise signals for the unperturbed and perturbed wavefields, respectively. Using the noise-contaminated signals, we define the noise-contaminated cross-correlation coefficient as

$$r'(t_s;t,t_w) = \frac{(u'_u,u'_p)_{(t_s;t,t_w)}}{\sqrt{(u'_u,u'_u)_{(t_s=0;t,t_w)}\,(u'_p,u'_p)_{(t_s=0;t,t_w)}}}\,. \tag{7.6}$$

Our aim is to derive a correction factor $c(t_s;t,t_w)$ such that

$$r(t_s;t,t_w) \approx c(t_s;t,t_w)r'(t_s;t,t_w)\,. \tag{7.7}$$

Throughout the remaining derivation, we assume that the noise signals $n_u(t)$ and $n_p(t)$ are realizations of mutually uncorrelated stationary stochastic processes with zero mean. Stationarity is here meant in the wide sense, i.e., the mean is the same (zero) for all times, and the autocorrelation depends only on the time shift t_s (Papoulis, 1991, p. 298).

Using equation (7.5), we find

$$(u'_u,u'_p)_{(t_s;t,t_w)} = (u_u,u_p)_{(t_s;t,t_w)} + (u_u,n_p)_{(t_s;t,t_w)} + (u_p,n_u)_{(t_s;t,t_w)} + (n_u,n_p)_{(t_s;t,t_w)}\,, \tag{7.8}$$

$$(u'_u,u'_u)_{(t_s=0;t,t_w)} = (u_u,u_u)_{(t_s=0;t,t_w)} + 2\,(u_u,n_u)_{(t_s=0;t,t_w)} + (n_u,n_u)_{(t_s=0;t,t_w)}\,, \tag{7.9}$$

$$(u'_p,u'_p)_{(t_s=0;t,t_w)} = (u_p,u_p)_{(t_s=0;t,t_w)} + 2\,(u_p,n_p)_{(t_s=0;t,t_w)} + (n_p,n_p)_{(t_s=0;t,t_w)}\,. \tag{7.10}$$

In equations (7.8)-(7.10), terms of the form $(u,n)_{(t_s;t,t_w)}$ appear, where $u = u(t)$ is a deterministic signal and $n = n(t)$ a realization of a stochastic process $\boldsymbol{n}(t)$[2]. These terms are time-windowed averages of the product $x(t) := u(t)n(t+t_s)$, where $u(t)$ is deterministic. Hence, $x(t)$ is a single realization of a stochastic process $\boldsymbol{x}(t)$ with the same statistical properties as $\boldsymbol{n}(t)$. Since we are dealing with windowed time-averages of single realizations of stochastic processes, we want to evaluate the time averages using the underlying ensemble averages. This leads us to the ergodicity of the stochastic signals treated here.

It is known (Papoulis, 1991, Chapter 13-1) that a (stationary) stochastic process $\boldsymbol{x}(t)$

[2]Throughout this paper we adopt the notation that $n(t)$ denotes a realization of an underlying stochastic process $\boldsymbol{n}(t)$.

is mean-ergodic if and only if $A(T) := \left(\int_0^T C_{\boldsymbol{xx}}(\tau)d\tau\right)/T \rightarrow 0$ as $T \rightarrow \infty$, where $C_{\boldsymbol{xx}}(\tau)$ is the autocovariance of $\boldsymbol{x}(t)$ at time-shift τ; this is Slutsky's theorem. For $A(T) \rightarrow 0$ as $T \rightarrow \infty$, it is sufficient for $C_{\boldsymbol{xx}}(\tau) \rightarrow 0$ as $\tau \rightarrow \infty$. This means that $\boldsymbol{x}(t)$ and $\boldsymbol{x}(t+\tau)$, and thus $\boldsymbol{n}(t)$ and $\boldsymbol{n}(t+\tau)$, must be uncorrelated for large time shifts τ, a requirement that is a reasonable assumption in most practical applications. Hence, with this assumption, we can conclude that $\boldsymbol{x}(t)$ is mean-ergodic, meaning that the time average $\left(\int_{-T}^{T} x(t)dt\right)/(2T)$, i.e., the time average of a single realization $x(t)$, is close to $E(\boldsymbol{x}(t)) = 0$ with probability close to 1. Here, $E(\boldsymbol{x}(t))$ denotes the expected value of $\boldsymbol{x}(t)$. We have $E(\boldsymbol{x}(t)) = 0$ since $\boldsymbol{x}(t)$ has the same statistical properties as $\boldsymbol{n}(t)$, i.e., zero mean. In turn, this means that

$$(u, n)_{(t_s;t,t_w)} = \frac{1}{2t_w}\int_{t-t_w}^{t+t_w} u(t')n(t'+t_s)dt' \approx 0\,, \tag{7.11}$$

provided the window length t_w is at least several dominant periods of the noise-contaminated signal. Note that the only real restrictive requirement used to show (7.11), is that $\boldsymbol{n}(t)$ is stationary with zero mean.

For the term $(n_u, n_p)_{(t_s;t,t_w)}$ in equations (7.8)-(7.10), we follow a similar argument, except that now the stochastic process to be considered is $\boldsymbol{z}(t) := \boldsymbol{n_u}(t)\boldsymbol{n_p}(t+t_s)$, where $\boldsymbol{n_u}(t)$ and $\boldsymbol{n_p}(t)$ are the underlying stochastic processes of which $n_u(t)$ and $n_p(t)$, respectively, are single realizations. As in the previous paragraph, for $\boldsymbol{z}(t)$ to be mean ergodic, it is sufficient if $C_{\boldsymbol{zz}}(\tau) \rightarrow 0$ as $\tau \rightarrow \infty$. Since $\boldsymbol{n_u}(t)$ and $\boldsymbol{n_p}(t)$ are assumed to be mutually uncorrelated both with zero mean, it follows that $C_{\boldsymbol{zz}}(\tau) = C_{\boldsymbol{n_u n_u}}(\tau)C_{\boldsymbol{n_p n_p}}(\tau)$. Again, under the often practical assumption that $\boldsymbol{n_u}(t)$ and $\boldsymbol{n_p}(t+\tau)$, as well as $\boldsymbol{n_p}(t)$ and $\boldsymbol{n_p}(t+\tau)$, are uncorrelated for large τ, we have that $C_{\boldsymbol{n_u n_u}}(\tau) \rightarrow 0$ and $C_{\boldsymbol{n_p n_p}}(\tau) \rightarrow 0$ as $\tau \rightarrow \infty$; hence $C_{\boldsymbol{zz}}(\tau) \rightarrow 0$ as $\tau \rightarrow \infty$. This means that $\boldsymbol{z}(t)$ is mean ergodic, which is equivalent to $\boldsymbol{n_u}(t)$ and $\boldsymbol{n_p}(t)$ being cross-covariance ergodic. Therefore, we have that $\left(\int_{-T}^{T} z(t)dt\right)/(2T)$ is close to $E(\boldsymbol{z}(t)) = E(\boldsymbol{n_u}(t)\boldsymbol{n_p}(t+t_s)) = E(\boldsymbol{n_u}(t))E(\boldsymbol{n_p}(t+t_s)) = 0$ with probability close to 1. This implies that we have

$$(n_u, n_p)_{(t_s;t,t_w)} = \frac{1}{2t_w}\int_{t-t_w}^{t+t_w} n_u(t')n_p(t'+t_s)dt' \approx 0\,, \tag{7.12}$$

with the same assumption for the window length t_w as used in equation (7.11). Note that both approximations (7.11) and (7.12) become more accurate with increasing window lengths.

Using equations (7.11) and (7.12) in equations (7.8)-(7.10), it follows that

$$\left(u'_u, u'_p\right)_{(t_s;t,t_w)} \approx (u_u, u_p)_{(t_s;t,t_w)}\,, \tag{7.13}$$

$$\left(u'_u, u'_u\right)_{(t_s=0;t,t_w)} \approx (u_u, u_u)_{(t_s=0;t,t_w)} + (n_u, n_u)_{(t_s=0;t,t_w)}\,, \tag{7.14}$$

$$\left(u'_p, u'_p\right)_{(t_s=0;t,t_w)} \approx (u_p, u_p)_{(t_s=0;t,t_w)} + (n_p, n_p)_{(t_s=0;t,t_w)}\,. \tag{7.15}$$

Using these approximations in equation (7.6), and substituting the resulting expression in

equation (7.7), it follows that the correction factor $c(t_s;t,t_w)$ is given by

$$c(t_s;t,t_w) \approx \frac{1}{\sqrt{\left(1-\frac{(n_u,n_u)_{(t_s=0;t,t_w)}}{(u'_u,u'_u)_{(t_s=0;t,t_w)}}\right)\left(1-\frac{(n_p,n_p)_{(t_s=0;t,t_w)}}{(u'_p,u'_p)_{(t_s=0;t,t_w)}}\right)}}. \tag{7.16}$$

This correction factor $c(t_s;t,t_w)$ depends on the unknown noise realizations $n_u(t)$ and $n_p(t)$. Note, that the correction factor contains only the zero-lag autocorrelations of these unknown noise realizations,and that therefore only an estimate of their respective variances is necessary to calculate the correction factor. In practice, assuming the noise is stationary, we can estimate these variances using the recorded wavefields $u'_u(t)$ and $u'_p(t)$ before $t=0$ of the experiment. Replacing $n_u(t)$ and $n_p(t)$ in equation (7.16) with the recorded wavefields before $t=0$ is a good approximation only if the terms of the form $(u,n)_{(t_s;t,t_w)}$ and $(n_u,n_p)_{(t_s;t,t_w)}$ in equations (7.8)-(7.10) are small compared to the remaining terms. Hence, for this replacement to be valid, we want the following inequalities to hold:

$$|(u_u+n_u,n_p)_{(t_s;t,t_w)}+(u_p,n_u)_{(t_s;t,t_w)}| \ll |(u_u,u_p)_{(t_s;t,t_w)}|, \tag{7.17}$$

$$2|(u_u,n_u)_{(t_s=0;t,t_w)}| \ll |(u_u,u_u)_{(t_s=0;t,t_w)}+(n_u,n_u)_{(t_s=0;t,t_w)}|, \tag{7.18}$$

$$2|(u_p,n_p)_{(t_s=0;t,t_w)}| \ll |(u_p,u_p)_{(t_s=0;t,t_w)}+(n_p,n_p)_{(t_s=0;t,t_w)}|, \tag{7.19}$$

where we have used the linearity of the $(\cdot,\cdot)_{(t_s;t,t_w)}$in equation (7.17). In Appendix O we rewrite these three inequalities into one inequality that can be evaluated using only the noise contaminated signals $u'_u(t)$ and $u'_p(t)$, and a representative noise realization $n_0(t)$, e.g., estimated from the wavefields before $t=0$ of the experiment. The resulting inequality [equation (O.5)] then determines if a certain time window has associated with it a reliable correction factor, when the unknown noise signals $n_u(t)$ and $n_p(t)$ in equation (7.16) are replaced with $n_0(t)$. Since for the applications treated in this work, (i.e., a displacement of the source and the displacement of the scatterer locations) the maximum correlation occurs at time-shift $t_s=\langle\tau\rangle=0$ s, we treat the special case $t_s=0$ in Appendix O only.

Replacing $n_u(t)$ and $n_p(t)$ with $n_0(t)$ in equation (7.16), the correction factor is given by

$$c(t_s;t,t_w) \approx \frac{1}{\sqrt{\left(1-\frac{(n_0,n_0)_{(t_s=0;t,t_w)}}{(u'_u,u'_u)_{(t_s=0;t,t_w)}}\right)\left(1-\frac{(n_0,n_0)_{(t_s=0;t,t_w)}}{(u'_p,u'_p)_{(t_s=0;t,t_w)}}\right)}}, \tag{7.20}$$

and is considered reliable if inequality (O.5) is satisfied with an appropriate value of γ [i.e., $O(10^{-1})$]. To derive condition (O.5), we assume that the noise is stationary and that $n_u(t)$ and $n_p(t)$ have about the same noise levels, (i.e., the same variance). If the noise levels of $n_u(t)$ and $n_p(t)$ differ substantially, separate estimates of the variances $(n_u,n_u)_{(t_s=0;t,t_w)}$ and $(n_p,n_p)_{(t_s=0;t,t_w)}$ can be used in equation (7.16) to calculate the correction factor. In this case condition (O.5) does not apply, and a new condition could be derived. To avoid

belaboring the point, we refrain from such a treatment.

7.5 Displacement of the scatterers

The problem of inferring the average displacement of scatterers in a strongly scattering medium from the multiply scattered wavefields has been used to study fluidized particle suspensions (Weitz & Pine, 1993; Heckmeier & Maret, 1997; Page *et al.*, 1999; Cowan *et al.*, 2000; Page *et al.*, 2000; Cowan *et al.*, 2002). In geophysics, this problem could be relevant when a strongly scattering region in the earth is strained, causing the scattering heterogeneities to move. In such a situation the displacement of the scatterers is not expected to be random, but will be correlated among scatterers. Here, we present a numerical experiment with point scatterers in a homogeneous background model, where we randomly perturb the scatterer locations and use CWI to infer their root-mean-square (RMS) displacement. Although this experiment is not directly related to a changing strain in the earth, it serves the purpose of testing the workings of our correction factor.

Snieder & Scales (1998) showed that for independent perturbations of the scatterer positions and isotropic scattering, the variance of the path length L is given by

$$\sigma_L^2 = 2n\delta^2 \,, \tag{7.21}$$

where n is the number of scatterers along the path, and δ is the RMS displacement of the scatterers in the direction of either coordinate axis (horizontal and vertical for two dimensions). Note that Snieder and Scales assume all directions of random displacement to be equally likely. As a result, the RMS displacement is the same in each direction (horizontal and vertical for 2D), i.e., for 2D the true RMS displacement would be $\sqrt{2}\delta$. Using that the number of scatterers is on average given by $n = vt/l^*$, with l^* the transport mean free path (Lagendijk & van Tiggelen, 1996), and v the velocity, and using $L = vt$, it follows that the variance of the traveltime perturbations is given by

$$\sigma_\tau^2 = \frac{2\delta^2 t}{vl^*} \,. \tag{7.22}$$

Inserting this in equation (7.2), it follows that the RMS displacement δ can be found from

$$\delta = \sqrt{(1 - \max\left[r(t_s; t, t_w)\right]) \frac{vl^*}{\omega^2 t}} \,. \tag{7.23}$$

Since the perturbations of the scatterer locations are assumed to be independent, and since the scattering is assumed to be isotropic, the mean traveltime perturbation $\langle \tau \rangle = 0$. This means that the maximum of the time-windowed cross-correlation function occurs at zero lag, i.e., $t_s = 0$.

Figure 7.1 shows the setup of our numerical experiment to test the inference of the scatterer displacements from the seismic coda using equation (7.23). This experiment was also outlined in Snieder *et al.* (2002). One hundred point scatterers (solid circles) are

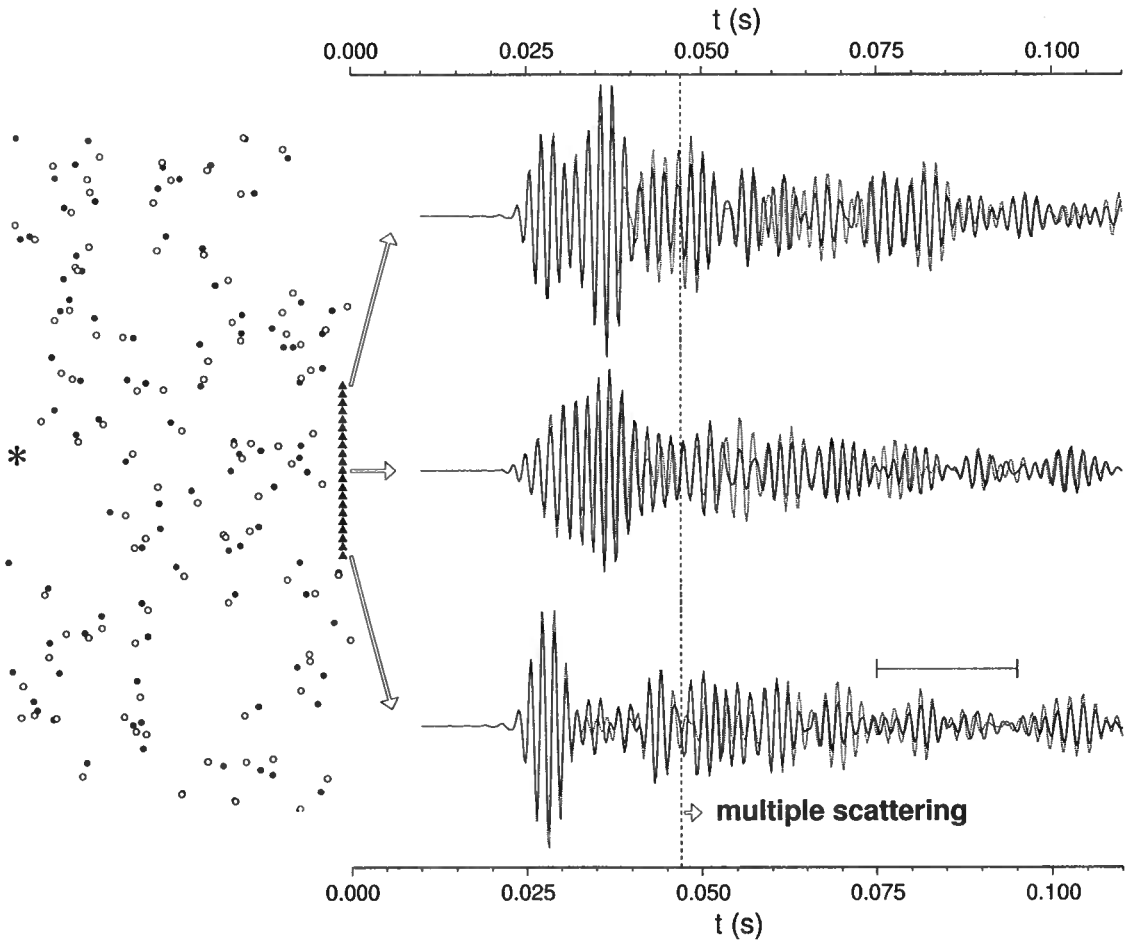


Figure 7.1. Unperturbed (filled circles) and perturbed (open circles) locations of the point scatterers in the medium. For display purposes the differences between the unperturbed and perturbed locations are magnified by a factor 10. The star denotes the source and the triangles the receivers. The modeled seismograms are shown for three receivers. The drawn seismograms are related to the unperturbed scatterer locations, and the dotted seismograms to the perturbed locations. For times later than $t = 4.7 \times 10^{-2}$ s (marked by the dashed line), the waves have scattered more than four times. The horizontal bar indicates the window length for the time-windowed cross-correlations used to calculate the RMS displacement δ shown in Figure 7.2.

contained in an area of $40 \times 80\,\mathrm{m}^2$, and the waveforms are calculated using a numerical implementation (Groenenboom & Snieder, 1995) of Foldy's method (Foldy, 1945). The resulting seismograms are shown by the solid lines for three locations on the edge of the area, and the source location is indicated by the asterisk. In these calculations the scattering amplitude was set to $-4i$, in order to get the maximum possible scattering strength as constrained by the optical theorem [Groenenboom & Snieder (1995); in their notation we used $\gamma = 4$]. The background velocity equaled 1500 m/s, and the source spectrum $S(\omega) = e^{-\omega^2/\omega_0^2}$, with ω the angular frequency, $\omega_0 = 2\pi f_0$, and $f_0 = 600$ Hz. The frequency band used was $400 - 800$ Hz, with a dominant frequency of about 500 Hz resutling from tapering on either side of the spectrum. Since in our experiment we have isotropic scattering, the transport mean free path equals the mean free path, i.e., $l^* = l$. The mean free path in our simulation was measured to be $l = 17.6$ m, which, using $n = vt/l$, can be used to infer that after $t = 4.7 \times 10^{-2}$ s the waves are on average scattered more than four times. This time is indicated in Figure 7.1 by the dotted vertical line. The perturbed scatterer locations are indicated by the open circles in Figure 7.1. For display purposes the displacements are magnified by a factor 10. The actual RMS displacement is $\delta_{\mathrm{true}} = 8 \times 10^{-2}$ m in both the horizontal and vertical direction. This displacement equals just 1/38 of the dominant wavelength (the dominant wavelength $\lambda = 3$ m). The resulting waveforms calculated using the displaced scatterers are shown for three receivers by the dotted lines.

Figure 7.2a shows the inferred value of δ [using equation (7.23)] as a function of the central window time t, where the estimated values for δ from all 21 receivers were averaged. [The receiver spacing was chosen such that the calculated multiply scattered waveforms were uncorrelated, meaning they can be treated as independent.] The dotted lines show the average inferred value of δ as a function of time plus or minus one standard deviation. The half-window duration t_w was 10^{-2} s, resulting in a window length of 10 dominant periods, and $\sqrt{\overline{\omega^2}} = 3.66 \times 10^3$ rad/s. The vertical dashed lines indicated the range of validity of equation (7.23). For early times (i.e., $t < 4.7 \times 10^{-2}$ s) the relation $n = vt/l$ for the number of scatterings used in the derivation of equation (7.23) is not valid, and for late times (i.e., $t > 2.9 \times 10^{-1}$ s), the second-order Taylor approximation of the autocorrelation of the source signal is inaccurate by more than 15%. This latter time is indicated by the rightmost dotted vertical line in Figure 7.2. For reference, Figure 7.2b shows an unperturbed (black) and perturbed (red) wavefield from the central receiver seen in Figure 7.1; the inset shows a blow-up of the seismograms at late times, indicating that although the signal is small, it is well above numerical precision.

Within the range of validity of equation (7.23), Figure 7.2a shows that the true displacement is recovered within the range given by the average RMS displacement plus or minus one standard deviation. For late times, the correlation coefficient is close to zero and, according to equation (7.23), the inferred value of δ is then given by $\sqrt{vl/\left(\overline{\omega^2}t\right)}$. This function is indicated by the dashed line in Figure 7.2 and agrees well with the inferred value of δ for late times. For these times the inferred value of δ is of course no longer a good estimate of the true RMS displacement.

As mentioned in the introduction, when the wavefields are contaminated with noise, the

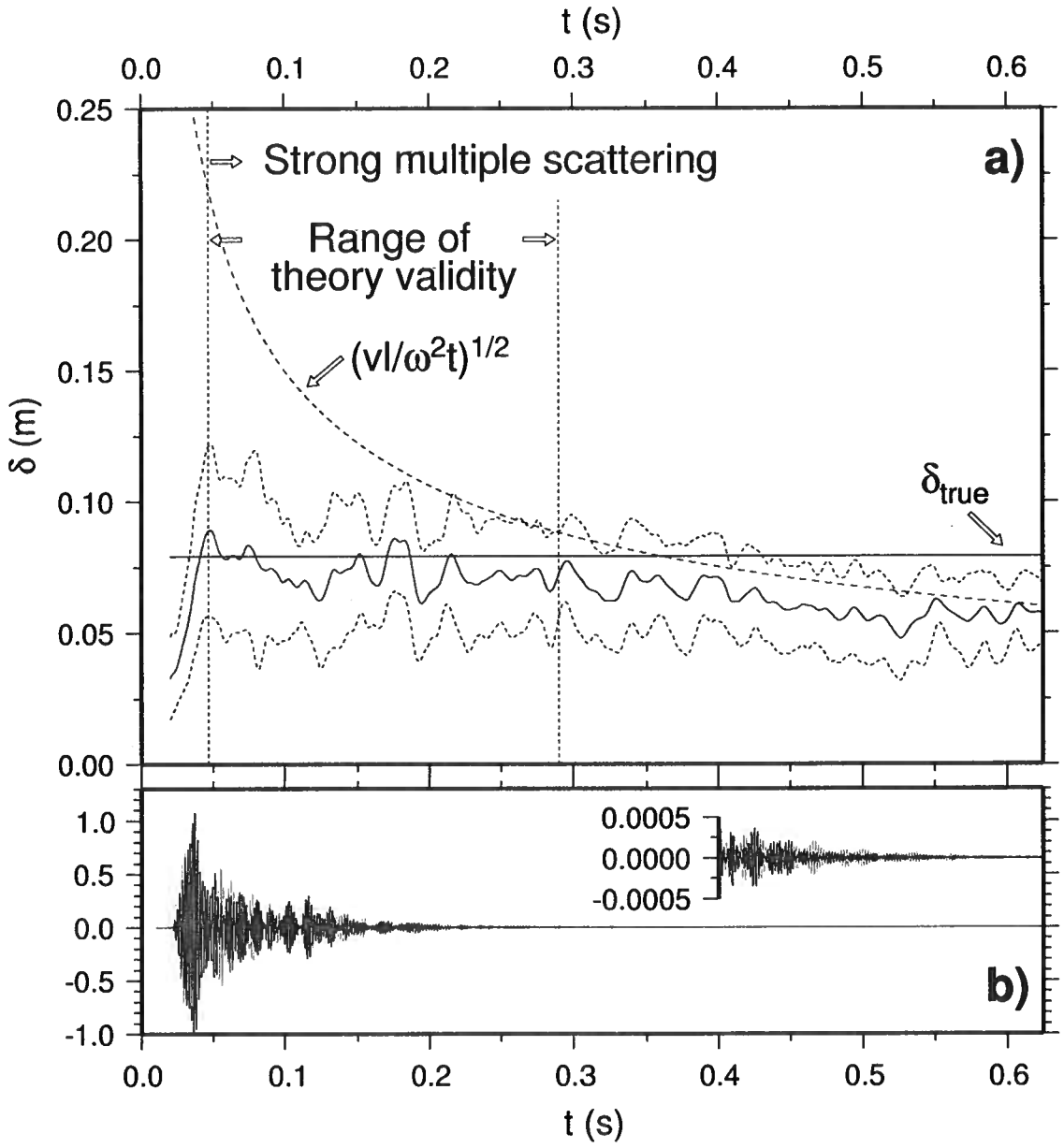


Figure 7.2. a) Inferred RMS displacement δ as a function of the center window time t (solid line) plus or minus one standard deviation (dotted lines). The true value of δ is indicated by the horizontal solid line. The range of validity of CWI is indicated by the vertical dashed lines. The half-window time t_w used in calculating the time windowed correlation coefficient equals 0.01s. b) Unperturbed (black) and perturbed (grey) seismogram from the central receiver seen in Figure 7.1; the inset shows a blow-up of the seismograms for $t > 0.4$ s.

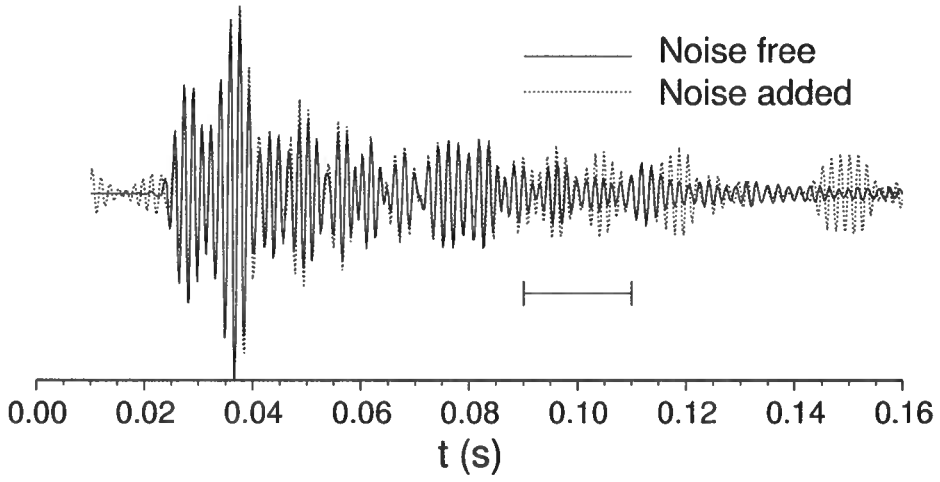


Figure 7.3. An example signal without noise (drawn) and with noise (dotted). The horizontal bar indicates the window length for the time-windowed cross-correlations used to calculate the RMS displacement δ shown in Figure 7.4.

number of independent time windows that can be used is limited to traveltimes where the ambient noise level is small compared to the amplitudes of the multiply scattered waves. In order to be able to use as many independent time windows as possible, and hence reduce the error bars on the inferred perturbation, it is important to correct the cross-correlation function for the bias caused by the noise. To assess the correction factor in equation (7.20), we added band-limited noise to the waveforms for all 21 receivers from our numerical experiment. The bandwidth of the noise was the same as that of the noise-free signals (i.e., $400-800$ Hz). Figure 7.3 shows a waveform with and without the added noise. Using the noise-contaminated waveforms, we again calculated the inferred values of δ both with and without the correction factor; see Figure 7.4c and 7.4b respectively. For reference, Figure 7.4a shows the inferred value of δ when the noise-free signals were used. Figure 7.4b shows that the noise induces a bias in the estimated value of δ; the presence of noise reduces the correlation between the unperturbed and perturbed waveforms, and causes the inferred values of δ, calculated using equation (7.23), to be larger. For early times, the true value of δ is embedded within the average value of δ plus or minus one standard deviation, but the estimated average value of δ is too high, especially for later times, where the lower amplitude values of the coda result in lower signal-to-noise ratios. Figure 7.4c shows that the correction factor from equation (7.20) accurately accounts for the bias caused by the noise. The noise estimates were obtained using the signal before the main first arrival in the seismograms. In the calculation of the results shown in Figure 7.4c, we used condition (O.5) for each receiver, with $\gamma = 0.125$ to select the time windows used for the inversion of δ [note that this is a different γ than that used by Groenenboom & Snieder (1995),

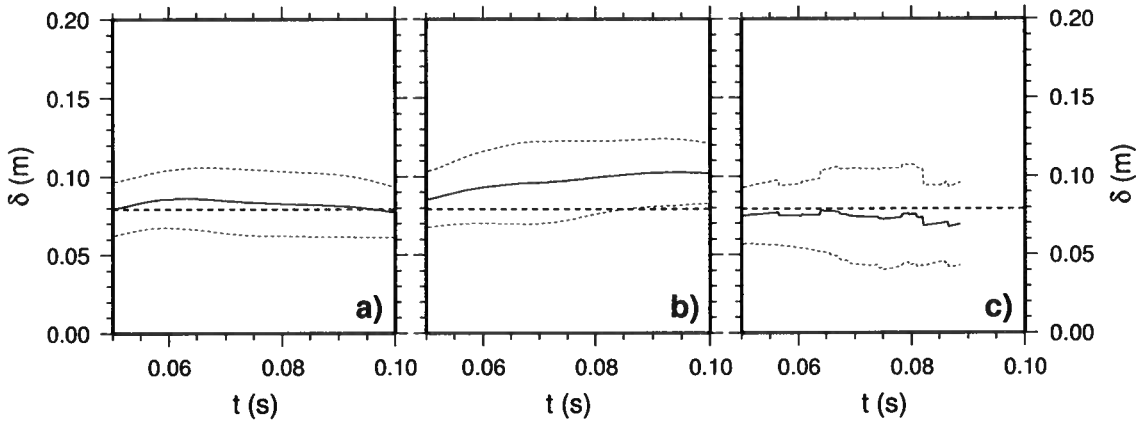


Figure 7.4. Inferred value of δ as a function of the central window time t, obtained using equation (7.23) with (a) the noise-free signals, (b) the noise-contaminated signals but no application of the correction factor, and (c) the noise-contaminated signals with application of the correction factor using equation (7.20).

where γ determines the scattering strength]. As a result, the number of receivers that had usable time windows for a given central window time t, varies for different t. This causes the jagged appearance of the average inferred value of δ (and the standard deviation). We used a time window only when at least seven receivers (i.e., 30% of the receivers) satisfied condition (O.5) at the central window time t. After $t \approx 9 \times 10^{-2}$ s fewer than seven time windows satisfied condition (O.5) with $\gamma = 0.125$, and hence the correction factor was judged unreliable. As a result, the noise-corrected estimates of the RMS displacement are not shown for times larger than $t \approx 9 \times 10^{-2}$ s.

7.6 Source separation

Snieder & Vrijlandt (2005) used CWI to estimate the distance between seismic events recorded at a single station, that have the same source mechanism. They derive the imprint of a change in source location on the variance of the traveltime differences, and then use equation (7.2) to infer this change from the maximum of the cross-correlation function, i.e., $\max[r(t_s; t, t_w)]$. In addition they show that for a change in source location this maximum occurs at time shift $t_s = \langle \tau \rangle = 0$, and that for two double-couple sources with a source separation in the fault plane, the relation between the source displacement Δs and the variance of the traveltimes σ_τ^2 is given by

$$\sigma_\tau^2 = \frac{\left(\frac{6}{\alpha^8} + \frac{7}{\beta^8}\right)}{7\left(\frac{2}{\alpha^6} + \frac{3}{\beta^6}\right)} (\Delta s)^2 , \tag{7.24}$$

where α and β are the P and S wave velocities, respectively. Different non-overlapping time-windows provide independent estimates of the source separation. These independent estimates in turn allow for calculation of error estimates of the source separation.

Figure 7.5a shows two seismograms (events 242003 and 242020) from earthquakes on the Hayward fault, California (Waldhauser & Ellsworth, 2000), recorded at station CSP of the Northern California Seismic Network. The recorded signal before the arrival of the P wave shows that the noise level is considerable. Figure 7.5b shows the maximum of the time-windowed cross-correlation function without the correction factor (thin line) and with the correction factor (thick line) applied, and Figure 7.5c shows the inferred values of the source separation using equation (7.24). Here the noise level (or variance) was estimated from the waveforms before the first arrivals. The half-window duration t_w used is 5 s (the full window length is indicated by the horizontal bar in Figure 7.5a), and the P and S wave velocities used to calculate the source displacement are $\alpha = 5750$ m/s and $\beta = 3320$ m/s, respectively. Note that we used overlapping time windows, since we plot $\max[r(t_s; t, t_w)]$ simply as a continuous function of the central window time t. Of course, non-overlapping windows could be used to ensure independent estimates of the source separation.

Figure 7.5b shows that the corrected values of $\max[r(t_s; t, t_w)]$ maintain a fairly constant level for times t late in the coda, whereas the uncorrected values decrease earlier in the coda because the noise decreases the similarity between both waveforms. As a result, the inferred values of the source separation using equation (7.24) are more or less constant even for larger traveltimes when the corrected values of $\max[r(t_s; t, t_w)]$ are used (Figure 7.5c). This indicates that the correction factor $c(t_s; t, t_w)$ given by equation (7.20), accurately corrects for the influence of the noise on the cross-correlation function. For very large times (say $t > 40$ s) the corrected values are more variable because the correction factor becomes unreliable. We purposely showed the times where the corrected values become variable, to indicate the level of variation caused by an unreliable correction factor. The time where the correction factor becomes unreliable could have been estimated using condition (O.5) with an appropriate value of γ [i.e., $O(10^{-1})$].

7.7 Conclusion

We have derived a correction factor for the influence of noise on the cross-correlation function, and have shown its accuracy using both numerical and field data. The derivation of the correction factor is based on the assumption that the stochastic processes underlying the noise realizations in both the unperturbed and perturbed signal are mutually uncorrelated and stationary with zero mean. We show the application of this correction factor in the context of CWI, for the inference of the RMS displacement of scatterers and a displacement of the source, from multiply scattered wavefields. For the displacement of the scatterers, we showed that in the presence of noise, a displacement of only 1/38 from the dominant wavelength can be successfully retrieved from the cross-correlation between the unperturbed and perturbed signals. This shows the power of CWI when compared to methods that use singly-scattered waves only.

Both for a displacement of the scatterer locations and for a change in the source

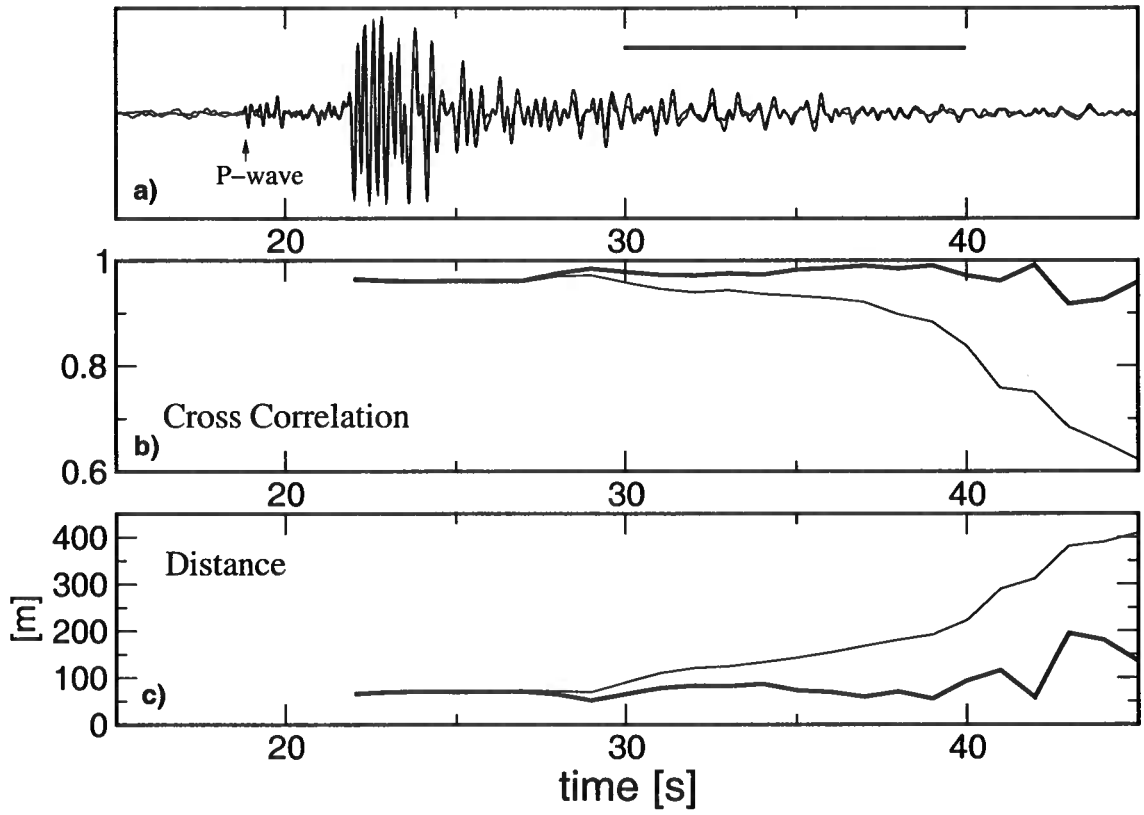


Figure 7.5. Seismograms from two earthquakes on the Hayward fault, California, recorded at station CSP of the Northern California Seismic Network (a), their cross-correlation maximum $\max[r(t_s; t, t_w)]$ as a function of the central window time t (b), uncorrected (thin line) and corrected (thick line), and the inferred source displacement (c) using both the uncorrected (thin line) and corrected (thick line) values of $\max[r(t_s; t, t_w)]$ shown in (b). The horizontal line in (a) indicates the window length used in the cross-correlation.

location, the perturbation is independent of the traveltimes of the multiply scattered waves. Using non-overlapping time windows to estimate the perturbations provides a consistency check of the method, and allows the calculation of error estimates. Therefore, our correction factor is relevant as it increases the number of usable time windows and hence allows for a reduction of the error estimates. In addition, the correction factor adjusts for bias in the cross-correlation induced by the noise. Since our factor depends on an estimate of the noise level in the data, we present a condition that allows determination of the reliability of the correction for a zero-lag cross-correlation. This condition can be verified using only noise-contaminated signals and an estimate of the noise level in the data. Using this condition, the time windows used in the zero-lag time-windowed cross-correlation can be judged to be reliable or not.

The use of the proposed correction factor is of course not limited to CWI. Any application that uses cross-correlations between different and noisy signals, and needs to correct for bias induced by noise, can benefit from the correction factor presented here.

7.8 Acknowledgment

We thank Mark Vrijlandt for Figure 7.5.

Chapter 8

Future work

A thorough investigation into a particular subject usually leads to new research questions or opportunities for future work. This is the case also for the work presented in this thesis. Therefore I conclude this thesis with an outline of future research opportunities that result from the presented work. Considering that this thesis covers several different topics, these opportunities are treated in the same order as the chapters to which they are related.

It is possible to derive pre-stack (common-offset) map time-migration equations in isotropic media like the ones presented in Chapter 2, which instead of the velocity use the offset horizontal slowness p_h (Fomel, 2005). This idea dates back to the work of Ottolini (1983). With these equations, the need for conventional velocity picking in CMP gathers is replaced with the need to estimate the local slopes p_h in CMP gathers, while the velocity becomes a data attribute that can derived from the local slopes. Fomel (2005) estimates these local slopes using plane-wave destruction filters (Fomel, 2002). Alternatively, the local slopes in the data can be estimated by projecting the data onto the curvelet frame. In the context of pre-stack time-migration with curvelets, this would require 3D extensions of curvelets for 2D imaging since the local slopes both in a common-offset and common-midpoint gather need to be known. In the context of velocity analysis in CMP gathers, however, 2D curvelets would be sufficient; since each curvelet has a central slope associated with it, each curvelet coefficient also has a velocity associated with it. Hence, once an intelligent thresholding procedure has determined which curvelets to keep to allow an accurate reconstruction of a CMP gather, the velocity profile results from a mapping of the curvelet coefficients to the velocity. In this way, velocity estimation in essence becomes a denoising problem. Also, since each curvelet really has a small range of local slopes associated with it, each curvelet could be mapped to a range of velocities. This allows an estimation of the uncertainty in the estimated velocity caused by the finite frequency character of the data.

In Chapter 4 I showed that the leading-order approximation (i.e., leading order in angular frequency, horizontal wavenumber, and migrated location) to common-offset (CO) time-migration can be obtained through a simple transformation of the input coordinates of the curvelets in the data, combined with amplitude scaling. This transformation is calculated using CO map time-migration that uses the local slopes provided by the curvelet decomposition of the data. In order to allow pre-stack depth-migration, the key ingredients to the derivation presented in Chapter 4 should be applied to the case of heterogeneous (anisotropic) media. These key ingredients are a stationary phase evaluation of the integral with respect to the non-constant variable, and subsequent linearization of the phase of the resulting oscillatory integral followed by linearization around the map-migrated location of

the center of the curvelet. This would provide the leading-order approximation to pre-stack depth migration. As pointed out in Chapter 4, such analysis could be done for demigration (i.e., modeling) in heterogeneous media to allow an output-based seismic imaging algorithm. That is, demigrating a curvelet in the image and subsequently checking whether the projection of the demigrated curvelet onto the data exceeds a certain threshold, would determine whether or not the curvelet in question contributes to the image. Considering the redundancy of the curvelet frame, however, applying this procedure to the whole image domain rather than a particular part of the image only, as in a *target-oriented* migration, would in principle mean demigrating more curvelets than there are samples in the image. Therefore, such a procedure would likely need to be combined with an initial estimate of the image, to determine the likely locations and orientations of the curvelets constituting the image. Such estimates could be obtained using map migration. The issue to either construct an output- or input-based imaging algorithm using curvelets remains a topic for future investigation.

A fast implementation of the leading-order approximation to seismic imaging using curvelets could be obtained if the coordinate transformation constituting the leading-order approximation (presented for CO time-migration in Chapter 4) would be expressed in terms of curvelet coefficients. In this case, a mapping of curvelet indices and associated coefficients could possibly be obtained. Such a procedure would allow for a fast computation of the leading-order approximation to seismic imaging in the curvelet domain. This would require an explicit calculation of the projection of a curvelet, subjected to the presented coordinate transformation, onto the curvelet frame. This seems a route worthwhile pursuing since indeed a fast algorithm to calculate the leading-order approximation would be obtained. However, an algorithm outside the curvelet domain that makes use of the compact support of curvelets in the Fourier domain (i.e., a wedge) seems worth investigating also.

In Chapter 4, the sparse representation of the data with curvelets was obtained using hard thresholding of the coefficients based on visually inspecting the difference between the original data and the data reconstructed with only a certain percentage of the (largest) curvelet coefficients. This simplistic thresholding results from focussing on the sparse representation of the imaging operator rather than sparse representation of the data, as mentioned in Chapter 4. More sophisticated thresholding techniques, however, exist to obtain a sparse representation of the data. Such thresholding techniques usually involve minimizing a certain misfit of the data combined with a regularization term specific to the problem. It remains to be investigated what measure of misfit and what regularization term are most suitable to obtain a sparse representation of seismic field data.

In the field of global seismology, tomographic studies these days no longer assume that traveltime perturbations from a reference background model are due to perturbations along the ray connecting source and receiver only (propagating through the background model). Rather they acknowledge the finite-frequency character of the data, which means that a finite volume surrounding this ray is responsible for the observed traveltime perturbations (see Nolet *et al.* (2005) for an introduction)[1]. This is described by a sensitivity kernel that

[1]As a matter of fact, in 3D traveltimes are unaffected by the presence of a (point) scatterer on this ray,

is used in the inversion of the observed traveltime perturbations. This kernel is calculated based on a single-scattering assumption by convolving the Green function from the source to a scatterer with the Green function from the same scatterer to the source, with an added scattering coefficient in between. Therefore, calculation of these kernels involves accurate calculation of the Green functions. Being able to calculate kernels for every region in the earth thus involves the ability to accurately calculate the needed Green functions, a computation that is time-consuming and thus expensive. There are two possible ways to reduce the computational cost (Tony Dahlen, personal communication). First, one can try to speed up the calculation of the needed Green functions. In light of the work of Smith (1998) that we mentioned in the context of extending the time migration with curvelets, presented in Chapter 4, to depth migration, it seems worthwhile to investigate the use of curvelets for the calculation of the Green functions to improve computational efficiency. Another approach would be to realize that for global studies the background model is always the same [usually the Preliminary Reference Earth Model (PREM) (Dziewonski & Anderson, 1981)]. That is, for global inversions the Green functions are always calculated for the same background model. Hence, if all Green functions would be calculated once, a tabulated approach would allow the calculation of the kernels necessary for the inversion. In this case, the issue is efficient storage of the huge table containing the wavefields in the PREM model. A sparse representation of these Green functions, and thus this huge table, could possibly be obtained using curvelets.

In recent investigations into the anisotropic nature of the subsurface, special emphasis is put on azimuthal variations of the seismic velocity [e.g., Jenner (2001) and Grechka *et al.* (2005)]. These investigations often start with or involve the azimuthal variations in the nonhyperbolic moveout observed in a CMP gather [e.g., Vasconcelos & Tsvankin (2006)]. Considering the accuracy of the rational-interpolation based approach to nonhyperbolic moveout analysis presented in Chapters 5 and 6, it seems worth investigating the application of this new method to estimate azimuthal anisotropy in the subsurface, and compare its accuracy with that of existing methods. Even though the accuracy of the rational-interpolation-based method for layered media, presented in Chapter 6, was established for synthetic data, its accuracy and applicability need to be verified for field data. The applicability to field data is currently being tested and will be reported on in a future paper that will include the work presented in Chapter 6.

---

because the traveltime from the source to a point on the ray to the receiver equals the traveltime along the ray. This argument can be extended to a scatterer of finite size that is (much) smaller than a wavelength, considering the effect of wavefront healing. In case wavefront healing occurs prior to the arrival of the wavefront at the receiver, the traveltime perturbation due to a small scatterer on the ray connecting source and receiver is also zero.

References

Akbar, F.E., Sen, M.E., & Stoffa, P.L. 1996. Prestack plane-wave Kirchhoff migration in laterally varying media. *Geophysics*, **61**, 1068–1079.

Alkhalifah, T. 1996. *Seismic Processing in Transversely Isotropic Media.* Ph.D. thesis, Colorado School of Mines, Center for Wave Phenomena.

Alkhalifah, T. 1997. Velocity analysis using nonhyperbolic moveout in transversely isotropic media. *Geophysics*, **62**, 1839–1854.

Alkhalifah, T. 1998. Acoustic approximations for processing in transversely isotropic media. *Geophysics*, **63**, 623–631.

Alkhalifah, T. 2000. An acoustic wave equation for anisotropic media. *Geophysics*, **65**, 1239–1250.

Alkhalifah, T., & Larner, K. 1994. Migration error in transversely isotropic media. *Geophysics*, **59**, 1405–1418.

Alkhalifah, T., & Tsvankin, I. 1995. Velocity analysis for transversely isotropic media. *Geophysics*, **60**, 1550–1566.

Alkhalifah, T., Tsvankin, I., Larner, K., & Toldi, J. 1996. Velocity analysis and imaging in transversely isotropic media: methodology and a case study. *The Leading Edge*, **15**, 371–378.

Baker, G.A. Jr. 1975. *Essentials of Padé Approximations.* Academic Press, Inc., New York.

Baker, G.A. Jr., & Graves-Morris, P. 1981. *Padé Approximants, part I: Basic Theory.* Addison-Wesley.

Bender, Carl M., & Orszag, Steven A. 1978. *Advanced Mathematical Methods for Scientists and Engineers.* New York, NY: McGraw-Hill.

Berrut, J.-P., & Mittelmann, H. 1997. Matrices for the direct determination of the barycentric weights of rational interpolation. *J. Comput. Appl. Math.*, **78**, 355–370.

Berrut, J.-P., & Mittelmann, H. 2000. Rational interpolation through the optimal attachment of poles to the interpolating polynomial. *Numer. Algorithms*, **23**, 315–328.

Billette, F., & Lambaré, G. 1998. Velocity macro-model estimation from seismic reflection data by stereotomography. *Geophysical Journal International*, **135**, 671–690.

Billette, F., Le Bégat, S., Podvin, P., & Lambaré, G. 2003. Practical aspects and applications of 2D stereotomography. *Geophysics*, **68**, 1008–1021.

Biondi, B., & Palacharla, G. 1996. 3-D prestack migration of common-azimuth data. *Geophysics*, **61**(6), 1822–1832.

Bishop, T.N., Bube, K.P., Cutler, R.T., Langan, R.T., Love, P.L., Resnick, J.R., Shuey, R.T., Spindler, D.A., & Wyld, H.W. 1985. Tomographic determination of velocity and depth in laterally varying media. *Geophysics*, **50**, 903–923.

Bleistein, N., Cohen, J.K., & Stockwell, J.W.Jr. 2000. *Mathematics of Multidimensional Seismic Imaging, Migration, and Inversion.* Springer.

Boyer, C.B. 1991. *A history of mathematics.* 2nd edn. Wiley and Sons, Inc.

Brandsberg-Dahl, S., de Hoop, M.V., & Ursin, B. 2003. Focusing in dip and AVA compensation on scattering-angle/azimuth common image gathers. *Geophysics*, **68**(1), 232–254.

Candès, E. J., & Demanet, L. 2005. The curvelet representation of wave propagators is optimally sparse. *Communications on Pure and Applied Mathematics*, **58**, 1472–1528.

Candès, E. J., & Donoho, D. L. 2004a. DCTvUSFFT: Digital Curvelet Transform via Unequispaced Fast Fourier Transforms. *Technical Report, California Institute of Technology.*

Candès, E. J., & Donoho, D. L. 2004b. New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities. *Comm. on Pure and Appl. Math.*, **57**, 219–266.

Candès, E. J., & Guo, F. 2002. New multiscale transforms, minimum total variation synthesis: Applications to edge-preserving image reconstruction. *Signal Processing*, **82**, 1519–1543.

Candès, E. J., Demanet, L., Donoho, D., & Ying, L. 2005. Fast discrete curvelet transforms. *preprint.*

Candès, E.J., & Donoho., D.L. 2000. *Curvelets - a surprisingly effective nonadaptive representation for objects with edges.* Vanderbilt University Press. Pages 105–120.

Cardano, G. 1545. *Ars Magna.*

Chapman, C. 2004. *Fundamentals of Seismic Wave Propagation.* Cambridge University Press.

Claerbout, J. F. 1976. *Fundamentals of geophysical data processing with applications to petroleum prospecting.* McGraw Hill Company.

Claerbout, J. F. 1985. *Imaging the Earth's Interior.* Blackwell Scientific Publications.

Claerbout, J. F. 1992. *Earth soundings analysis: Processing versus inversion.* Blackwell Scientific Publications, Inc.

Claerbout, J.F. 1970. Coarse grid calculations of waves in inhomogeneous media with application to delineation of complicated seismic structure. *Geophysics*, **35**(3), 407–418.

Claerbout, J.F. 1971. Toward a unified theory of reflector mapping. *Geophysics*, **36**(3), 467–481.

Cohen, Jack K. 1998. A convenient expression for the NMO velocity function in terms of ray parameter. *Geophysics*, **63**, 275–278.

Coleman, T.F., & Li, Y. 1994. On the Convergence of Reflective Newton Methods for Large-Scale Nonlinear Minimization Subject to Bounds. *Mathematical Programming*, **67**, 189–224.

Coleman, T.F., & Li, Y. 1996. An Interior, Trust Region Approach for Nonlinear Minimization Subject to Bounds. *SIAM Journal on Optimization*, **6**, 418–445.

Córdoba, A., & Fefferman, C. 1978. Wave packets and Fourier Integral Operators. *Comm. PDE's*, **3**, 979–1005.

Cowan, M.L., Page, J.H., & Weitz, D.A. 2000. Velocity fluctuations in fluidized suspensions probed by ultrasonic correlation spectroscopy. *Phys. Rev. Lett.*, **85**, 453–456.

Cowan, M.L., Jones, I.P., Page, J.H., & Weitz, D.A. 2002. Diffusing Acoustic Wave Spectroscopy. *Phys. Rev. E*, **65**, 066605–1/11.

de Hoop, M. V., & Brandsberg-Dahl, S. 2000. Maslov asymptotic extension of generalized Radon transform inversion in anisotropic elastic media: A least-squares approach. *Inverse Problems*, **16**, 519–562.

Dellinger, J., Muir, F., & Karrenback, M. 1993. Anelliptic approximations for TI media. *J. Seis. Expl.*, **2**, 23–40.

Dennis, J.E. Jr. 1977. *Nonlinear Least Squares.* Academic Press. Pages 269–312.

Derode, A., Roux, P., & Fink, M. 1995. Robust acoustic time reversal with high-order multiple scattering. *Phys. Rev. Lett.*, **75**, 4206–4209.

Derode, A., Tourin, A., & Fink, M. 1999. Ultrasonic pulse compression with one-bit time reversal through multiple scattering. *J. Appl. Phys.*, **85**, 6343–6352.

Dix, C.H. 1955. Seismic velocities from surface measurements. *Geophysics*, **20**, 68–86.

Douma, H., & Calvert, A. 2006. Nonhyperbolic moveout analysis in VTI media using rational interpolation. Accepted for publication in *Geophysics.*

Douma, H., & de Hoop, M. V. 2006. Explicit expressions for pre-stack map time-migration in isotropic and VTI media and the applicability of map depth-migration in heterogeneous anisotropic media. *Geophysics*, **71**(1), S13–S28.

Douma, H., & de Hoop, M.V. 2004. Wave-character preserving pre-stack map migration using curvelets. *In: Expanded Abstracts 74th annual international meeting of the SEG.*

Douma, H., & de Hoop, M.V. 2005. On common-offset pre-stack time-migration with curvelets. *In: Expanded Abstracts 75th annual international meeting of the SEG.*

Dziewonski, A. M., & Anderson, D. L. 1981. Preliminary Reference Earth Model. *Physics of the Earth and Planetary Interiors*, **25**, 297–356.

Fefferman, C. 1973. A note on spherical summation multipliers. *Israel J. Math.*, **15**, 44–52.

Fink, M. 1997. Time-reversed Acoustics. *Physics Today*, **50**, 34–40.

Foldy, L.O. 1945. The Multiple scattering of waves. *Phys. Rev.*, **67**, 107–119.

Fomel, S. 2002. Applications of plane-wave destruction filters. *Geophysics*, **67**(6), 1946–1960.

Fomel, S. 2004. On anelliptic approximations for qP velocities in VTI media. *Geophysical Prospecting*, **52**, 247–259.

Fomel, S. 2005. Velocity-independent time-domain seismic imaging using local event slopes. *In: Expanded abstracts of the 75th annual international meeting of the SEG.*

Fowler, P.J. 2003. Practical VTI approximations: a systematic anatomy. *Journal of Applied Geophysics*, **54**, 347–367.

Gjoystdal, H., & Ursin, B. 1981. Inversion of reflection times in three dimensions. *Geophysics*, **46**, 972–983.

Graeser, E., Lode, W., & Pott, G. 1957. Representation of depth-contour maps of arbitrarily curved reflection horizons, including refraction of rays, three-dimensional case. *Geophysical Prospecting*, **5**, 135–141.

Gray, W.C., & Golden, J.E. 1983. Velocity determination in a complex earth. Soc. Expl. Geophys.

Grechka, V., & Tsvankin, I. 1998. Feasibility of nonhyperbolic moveout inversion in transversely isotropic media. *Geophysics*, **63**, 957–969.

Grechka, V., Pech, A., & Tsvankin, I. 2005. Parameter estimation in orthorhombic media using multicomponent wide-azimuth reflection data. *Geophysics*, **70**, D1–D8.

Grêt, A., Snieder, R., Aster, R.C., & Kyle, P.R. 2005. Monitoring rapid temporal changes in a volcano with coda wave interferometry. *Geophys. Res. Lett.*, **32**, L06304, 10.1029/2004GL021143.

Grêt, A., Snieder, R., & Scales, J. 2006. Time-lapse monitoring of rock properties with coda wave interferometry. *J. Geophys. Res.*, **111**, B03305, doi:10.1029/2004JB003354.

Groenenboom, J., & Snieder, R.K. 1995. Attenuation, dispersion, and anisotropy by multiple scattering of transmitted waves through distributions of scatterers. *J. Acoust. Soc. Am.*, **98**(6), 1–11.

Guillemin, V. 1985. On some results of Gel'fand in integral geometry. *Proc. Symposia in Pure Math.*, **43**, 149–155.

Haas, A.G., & Viallix, J.R. 1976. Krigeage applied to geophysics. *Geophysical Prospecting*, **24**, 49–69.

Hake, H. 1986. Slant stacking and its significance for anisotropy. *Geophys. Prosp.*, **34**, 595–608.

Hake, H., Helbig, K., & Mesdag, C.S. 1984. Three-term Taylor series for t^2 - x^2 curves over layered transversely isotropic ground. *Geophys. Prosp.*, **32**, 828–850.

Harlan, W., & Burridge, R. 1983. A tomographic velocity inversion for unstacked data. *Stanford Exploration Project report SEP-37, 1-7.*

Heckmeier, M., & Maret, G. 1997. Dark Speckle Imaging of Colloidal Suspensions in Multiple Light Scattering Media. *Progr. Colloid. Polym. Sci.*, **104**, 12–16.

Helbig, K. 1994. *Foundations of Anisotropy for Exploration Seismics.* Pergamon.

Hermont, A.J. 1979. Letter to the editor, re: Seismic controllable directional reception as practiced in the USSR. *Geophysics*, **44**, 1601–1602.

Hernández, E., & Weiss, G. 1996. *A first course on wavelets.* CRC Press, Inc.

Herrmann, F. 2003a. Optimal imaging with curvelets. *In: Expanded Abstracts 73th annual international meeting of the SEG.*

Herrmann, Felix J. 2003b. *Multifractional splines: application to seismic imaging.* Proceedings of SPIE Technical Conference on Wavelets: Applications in Signal and Image Processing X, Volume 5207. Pages 240–258.

Hua, B., & McMechan, G. A. 2001. Parsimonious 2D poststack Kirchhoff depth migration. *Geophysics*, **66**, 1497–1503.

Hua, B., & McMechan, G. A. 2003. Parsimonious 2D prestack Kirchhoff depth migration. *Geophysics*, **68**, 1043–1051.

Iversen, E., & Gjoystdal, H. 1996. Event-oriented velocity estimation based on prestack data in time or depth domain. *Geophysical Prospecting*, **44**, 643–686.

Iversen, E., Gjoystdal, H., & Hansen, J.O. 2000. Prestack map migration as an engine for parameter estimation in TI media. *Pages 1004–1007 of: Expanded Abstracts 70th annual international meeting of the SEG.*

Jannane, H., Beydoen, W., Crase, E., Cao, D., Koren, Z., Landa, E., Menses, M., Pica, A., Noble, M., Roeth, G., Singh, S., Snieder, R., Tarantola, A., Trezeguet, D., & Xie, M. 1989. Wavelengths of Earth Structures that Can Be Resolved From Seismic Reflection Data. *Geophysics*, **54**, 906–910.

Jaramillo, H., & Larner, K. 1995. Prestack migration error in transversely isotropic media. *CWP Research Report (CWP-185).*

Jenner, E. 2001. *Azimuthal anisotropy of 3-D compressional wave seismic data, Weyburn field, Saskatchewan, Canada.* Ph.D. thesis, Colorado School of Mines.

Jones, L.E.A., & Wang, H.F. 1981. Ultrasonic velocities in Cretaceous shales from the Williston Basin. *Geophysics*, **46**, 288–297.

Kleyn, A.H. 1977. On the migration of reflection time contour maps. *Geophysical Prospecting*, **25**, 125–140.

Kravtsov, Y.A. 1988. Rays and Caustics As Physical Objects. *Pages 227–348 of:* Wolf, E. (ed), *Prog. in Optics, XXVI.* Amsterdam: Elsevier.

Lagendijk, A., & van Tiggelen, B.A. 1996. Resonant Multiple Scattering of Light. *Phys. Rep.*, **270**, 143–215.

Larner, K., & Cohen, J. 1993. Migration error in factorized transversely isotropic media with linear velocity variation with depth. *Geophysics*, **58**, 1454–1467.

Le Rousseau, J.H, & de Hoop, M.V. 2001. Modeling and imaging with the scalar generalized-screen algorithms in isotropic media. *Geophysics*, **66**(5), 1551–1568.

Luo, Y., & Schuster, G.T. 1991. Wave-equation traveltime tomography. *Geophysics*, **56**, 645–653.

Maher, S.M., & Hadley, D.M. 1985. Development of an accurate, stable, and interactive map migration algorithm. *Pages 551–552 of: Expanded abstracts of the 55th annual international meeting of the SEG, session: S15.8.*

Maher, S.M., Thorson, J.R., Hadley, D.M., & Swanger, H.J. 1987. Study of comparative interval velocities for map migration. *Pages 471–473 of: Expanded abstracts of the 57th Ann. Internat. Mtg. Soc. of Expl. Geophys, Session: S3.7.*

Mallat, S. G. 1998. *A wavelet tour of signal processing.* First edn. Academic Press, USA.

Maret, G., & Wolf, G.E. 1987. Effect of Brownian Motion of Scatterers. *Z. Phys. B*, **65**, 409–413.

Marquering, H., Dahlen, F.A., & Nolet, G. 1999. Three-dimensional sensitivity kernels for finite-frequency traveltimes: the banana-doughnut paradox. *Geophys. J. Int.*, **137**, 805–815.

Moré, J.J. 1977. *The Levenberg-Marquardt algorithm: implementation and theory.* Vol. 630. Springer Verlag. Pages 105–116.

Musgrave, A.W. 1961. Wave-front charts and three dimensional migrations. *Geophysics*, **26**, 738–753.

Nolet, G., Dahlen, T.A., & Montelli, R. 2005. Traveltimes and amplitudes of seismic waves: a re-assessment. *Pages 37–47 of:* Levander, A., & Nolet, G. (eds), *Seismic Earth.* American Geophysical Union.

Ottolini, R. 1983. Velocity independent seismic imaging. *Stanford Exploration Project report SEP-37, 1-7.*

Page, J.H., Jones, I.P., Schriemer, H.P., Cowan, M.L., Sheng, P., & Weitz, D.A. 1999. Diffusive transport of acoustic waves in strongly scattering media. *Physica B*, **263-264**, 37.

Page, J.H., Cowan, M.L., & Weitz, D.A. 2000. Diffusing acoustic wave spectroscopy of fluidized suspensions. *Physica B*, **279**, 130–133.

Papoulis, A. 1991. *Probability, Random Variables, and Stochastic Processes.* McGraw-Hill Inc.

Pine, D.J., Weitz, D.A., Chaikin, P.M., & Herbolzheimer, E. 1988. Diffusing Wave Spectroscopy. *Phys. Rev. Lett*, **60**, 1134–1137.

Reilly, J.M. 1991. Integrated interpretation, 3D map migration and VSP modelling project, northern U.K. southern gas basin. *Geophysical Prospecting*, **39**, 253–278.

Riabinkin, L.A. 1991. *Fundamentals of resolving power of controlled directional reception (CDR) of seismic waves.* Society of Exploration Geophysicists (Tulsa), Geophysics Reprint Series. Pages 36–60.

Richardson, M. 1958. *College Algebra.* Prentice Hall, Inc.

Rieber, F. 1936. A new reflection system with controlled directional sensitivity. *Geophysics*, **1**, 97–106.

Rieber, F. 1937. Complex reflection patters and their geological sources. *Geophysics*, **2**, 132–160.

Rietveld, W.E.A., & Berkhout, A.J. 1992. Depth migration combined with controlled illumination. *SEG Technical Program Expanded Abstracts*, **11**, 931–934.

Roberts, P.M., Scott Phillips, W., & Fehler, M.C. 1992. Development of the active doublet method for measuring small velocity and attenuation changes in solids. *J. Acoust. Soc. Am.*, **91**, 3291–3302.

Sarkar, D., Baumel, R.T., & Larner, K.L. 2002. Velocity analysis in the presence of amplitude variation. *Geophysics*, **67**(5), 1664–1672.

Sattlegger, J. 1964. Series for three-dimensional migration in reflection seismic interpretation. *Geophysical Prospecting*, **12**, 115–134.

Sattlegger, J.W., Stiller, P.K., Echterhoff, J.A., & Hentschke, M.K. 1980. Common offset plane migration. *Geophysical Prospecting*, **28**, 859–871.

Sayers, C.M. 1994. The elastic anisotropy of shales. *J. Geophys. Res.*, **99 (B1)**, 767–774.

Schoenberg, M. A., & de Hoop, M.V. 2000. Approximate dispersion relations for qP-qSV waves in transversely isotropic media. *Geophysics*, **65**, 919–933.

Smith, H. 1997. A Hardy space for Fourier integral operators. *J. Geom. Anal.*, **7**.

Smith, H. 1998. A parametrix construction for wave equations with $C^{1,1}$ coefficients. *Ann. Inst. Fourier*, **48**, 797–835.

Snieder, R. 2004. Coda wave interferometry. *Pages 54–56 of: 2004 McGraw-Hill yearbook of science & technology.* New York: McGraw-Hill.

Snieder, R. 2006. The theory of coda wave interferometry. *Pure Appl. Geophys., in press.*

Snieder, R., & Vrijlandt, M. 2005. Constraining Relative Source Locations with Coda Wave Interferometry: Theory and Application to Earthquake Doublets in the Hayward Fault, California. *J. Geophys. Res.*, **110**, B04301, 10.1029/2004JB003317.

Snieder, R.K. 2002. Coda wave interferometry and the equilibration of energy in elastic media. *Phys. Rev. E*, **66**, 046615–1,8.

Snieder, R.K., & Scales, J.A. 1998. Time reversed imaging as a diagnostic of wave and particle chaos. *Phys. Rev. E*, **58**, 5668–5675.

Snieder, R.K., Grêt, A., Douma, H., & Scales, J. 2002. Coda Wave Interferometry for estimating nonlinear behavior in seismic velocity. *Science*, **295**(22 March), 2253–2255.

Spetzler, J., & Snieder, R. 2004. The Fresnel volume and transmitted waves. *Geophysics*, **69**, 653–663.

Stein, E. M. 1993. *Harmonic Analysis: Real-variable methods, orthogonality, and oscillatory integrals.* Second edn. Princeton University Press.

Stoer, J., & Bulirsch, R. 1993. *Introduction to Numerical Analysis.* 2nd edn. Springer-Verlag.

Stoffa, P.L., Fokkema, J.T., de Luna Freire, R.M., & Kessinger, W.P. 1990. Split-step Fourier migration. *Geophysics*, **55**(4), 410–421.

Stolk, C. C. 2000. *On the modeling and inversion of seismic data.* Ph.D. thesis, Utrecht University.

Stolk, C. C., & De Hoop, M. V. 2002a. Microlocal analysis of seismic inverse scattering in anisotropic, elastic media. *Comm. Pure Appl. Math.*, **55**, 261–301.

Stolk, C. C., & De Hoop, M. V. 2002b. Seismic inverse scattering in the 'wave equation' approach. Submitted to *SIAM J. Appl. Math.*

Stovas, A., & Ursin, B. 2004. New travel-time approximations for a transversely isotropic medium. *J. Geophys. Eng.*, **1**, 128–133.

Strang, G., & Nguyen, T. 1997. *Wavelets and filter banks.* Revised edn. Wellesley-Cambridge Press.

Sword, C. H. 1987. *Tomographic determination of interval velocities from reflection seismic data: the method of controlled directional reception.* Ph.D. thesis, Stanford Exploration Project, Stanford University.

ten Kroode, A.P.E., Smit, D.-J., & Verdel, A. R. 1998. A microlocal analysis of migration. *Wave Motion*, **28**, 149–172.

Thomsen, L. 1986. Weak elastic anisotropy. *Geophysics*, **51**, 1954–1966.

Toldi, J., Alkhalifah, T., Berthet, P., Arnaud, J., Williamson, P., & Conche, B. 1999. Case study of estimation of anisotropy. *The Leading Edge*, **18**, 588–593.

Tsvankin, I. 1996. P-wave signatures and notation for transversely isotropic media: An overview. *Geophysics*, **61**, 467–483.

Tsvankin, I. 2001. *Seismic signatures and analysis of reflection data in anisotropic media.* Elsevier Science Ltd.

Tsvankin, I., & Thomsen, L. 1994. Nonhyperbolic reflection moveout in anisotropic media. *Geophysics*, **59**, 1290–1304.

Ursin, B., & Stovas, A. 2006. Traveltime approximations for a transversely isotropic medium. *accepted for publication in Geophysics.*

Van der Baan, M. 2004. Processing of anisotropic data in the $\tau - p$ domain: I-Geometric spreading and moveout corrections. *Geophysics*, **69**, 719–730.

Van der Baan, M., & Kendall, J. M. 2002. Estimating anisotropy parameters and travel-times in the τ-p domain. *Geophysics*, **67**, 1076–1086.

van Trier, J.A. 1990. *Tomographic determination of structural velocities from depth-migrated seismic data.* Ph.D. thesis, Stanford Exploration Project, Stanford University.

Vasconcelos, I., & Tsvankin, I. 2006. Nonhyperbolic moveout inversion of wide-azimuth P-wave data for orthorhombic media. *Geophysical Prospecting*, in press.

Červený, V. 2001. *Seismic Ray Theory.* Cambridge University Press.

Vlaar, N.J. 1968. Ray theory for an anisotropic inhomogeneous elastic medium. *Bulletin of the Seismological Society of America*, **58**, 2053–2072.

Waldhauser, F., & Ellsworth, W. 2000. A double-difference earthquake location algorithm: Method and application to the northern Hayward fault, California. *Bull. Seism. Soc. Am.*, **90**, 1353–1368.

Weber, M. 1955. Die bestimmung einer beliebig gekruemmten schichtgrenze aus seismischen reflexionsmessungen. *Geofisica pura e applicata*, **32**, 7–11.

Weitz, D.A., & Pine, D.J. 1993. Diffusing Wave Spectroscopy. *Page 652 of:* Brown, W. (ed), *Dynamic Light Scattering.* Clarendon Press, Oxford.

Whitcombe, D.N., & Carroll, R.J. 1994. The application of map migration to 2-D migrated data. *Geophysics*, **59**, 1121–1132.

Winterstein, D.F. 1986. Anisotropy effects in P-wave and SH-wave stacking velocities contain information in lithology. *Geophysics*, **51**, 661–672.

Wookey, J., Van der Baan, M., Smit, D., & Kendall, J.-M. 2002. Tau-p domain VTI parameter inversions using limited-offset data. *Page F041 of: Extended Abstracts, 64th Meeting, European Association of Geoscientists and Engineers.*

Ying, L, Demanet, L., & Candés, E. 2005. 3D discrete curvelet transform. *preprint, available at http://www.curvelet.org.*

Yodh, A.G., Kaplan, P.D., & Pine, D.J. 1990. Pulsed diffusing-wave spectroscopy: High resolution through nonlinear optical gating. *Phys. Rev. B*, **42**, 4744–4747.

Zavalishin, B.R. 1981. Perfection of methods for constructing seismic images using controlled directional reception. *Soviet Geology and Geophysics*, **22**, 98–104.

Zhang, F., & Uren, N. 2001. Approximate explicit ray velocity functions and travel times for P-waves in TI media. *Pages 106–109 of: Expanded abstracts of the 71st Ann. Internat. Mtg. Soc. of Expl. Geophys, San Antonio, Texas, USA.*

Appendix A

From phase angle to horizontal slowness assuming $V_{S0} = 0$

The slowness surface for qP waves is convex, which in combination with a VTI medium, assures that the only branch points in vertical slowness occur at $\theta = \pm\pi/2$, where θ is the phase angle measured from the vertical axis of symmetry. In homogeneous media, given the surface seismic geometry, these branch points are never reached, since turning waves do not occur in such media. Thus, for qP waves in such media we can parametrize the phase velocity and its derivative uniquely in terms of the horizontal slowness p.

Substituting $\sin\theta = pV$ in equation (2.35) and solving for V, leads to

$$V(p) = V_{P0}\sqrt{\frac{1 + 2p^2V_{P0}^2(\delta - \epsilon)}{1 - 2p^2V_{P0}^2\epsilon + 2p^4V_{P0}^4(\delta - \epsilon)}}\,, \tag{A.1}$$

where we used that $V_{S0} = 0$ (i.e., $f = 1$) and that the kinematics of qP-waves in anisotropic media are independent of V_{S0} within the limits of seismic accuracy Alkhalifah (1998). The derivative $\frac{dV}{dp}$ is then

$$\frac{dV}{dp} = \frac{2pV_{P0}^3\delta - 4p^3V_{P0}^5(\delta - \epsilon) - 4p^5V_{P0}^7(\delta - \epsilon)^2}{\left(1 - 2p^2V_{P0}^2\epsilon + 2p^4V_{P0}^4(\delta - \epsilon)\right)^{3/2}\sqrt{1 + 2p^2V_{P0}^2(\delta - \epsilon)}}\,. \tag{A.2}$$

Both expressions could as well be readily rewritten in terms of η, $V_{NMO}(0)$, and V_{P0}.

Using these expressions, we can write the vertical slowness $q = \sqrt{1/V^2 - p^2}$ as

$$q = \frac{1}{V_{P0}}\sqrt{\frac{1 - p^2V_{P0}^2(1 + 2\epsilon)}{1 + 2p^2V_{P0}^2(\delta - \epsilon)}}\,, \tag{A.3}$$

or in terms of η, $V_{NMO}(0)$, and V_{P0} [see also Alkhalifah (1998, equation A-10)],

$$q = \frac{1}{V_{P0}}\sqrt{1 - \frac{p^2V_{NMO}^2(0)}{1 - 2p^2V_{NMO}^2(0)\eta}}\,, \tag{A.4}$$

In order to calculate the group angle [see equation (2.33)] as a function of the horizontal

slowness, we use the chain rule,

$$\frac{dV}{d\theta} = \frac{dV}{dp}\frac{dp}{d\theta} . \tag{A.5}$$

Using $p = \sin\theta/V$, this becomes

$$\frac{dV}{d\theta} = \left(\frac{dV}{dp}\sqrt{1-p^2V^2}\right) / \left(V + p\frac{dV}{dp}\right) , \tag{A.6}$$

where we used that in homogeneous media we have $\cos(\theta) = \sqrt{1-p^2V^2} > 0$. Then, using this expression together with equations (A.1) and (A.2) in equation (2.33), we find

$$\tan\psi = \frac{pV_{P0}\left(1+2\delta\right)}{\left(1+2p^2V_{P0}^2\left(\delta-\epsilon\right)\right)^{3/2}\sqrt{1-p^2V_{P0}^2\left(1+2\epsilon\right)}} , \tag{A.7}$$

or in terms of η, $V_{NMO}(0)$ and V_{P0},

$$\tan\psi = \frac{pV_{NMO}^2(0)}{V_{P0}\left(1-2p^2V_{NMO}^2(0)\eta\right)^{3/2}\sqrt{1-p^2V_{NMO}^2(0)\left(1+2\eta\right)}} . \tag{A.8}$$

Using the simplified expressions (A.4) and (A.8) for the group angle and vertical slowness in equations (2.61) and (2.63)-(2.65) then leads to the following system of equations for pre-stack map demigration in homogeneous VTI media,

$$\begin{aligned} 0 &= \frac{p_s\sqrt{1-s_\gamma^2}}{\left(1-2\eta p_s^2V_{NMO}^2(0)\right)^{3/2}\sqrt{1-p_s^2V_{NMO}^2(0)\left(1+2\eta\right)}} \\ &\quad - \frac{p_r\sqrt{1-r_\gamma^2}}{\left(1-2\eta p_r^2V_{NMO}^2(0)\right)^{3/2}\sqrt{1-p_r^2V_{NMO}^2(0)\left(1+2\eta\right)}} , \end{aligned} \tag{A.9}$$

$$\begin{aligned} \frac{4h}{V_{NMO}^2(0)} &= \frac{t_m r_\gamma p_r}{\left(1-2\eta p_r^2V_{NMO}^2(0)\right)^{3/2}\sqrt{1-p_r^2V_{NMO}^2(0)\left(1+2\eta\right)}} \\ &\quad - \frac{t_m s_\gamma p_s}{\left(1-2\eta p_s^2V_{NMO}^2(0)\right)^{3/2}\sqrt{1-p_s^2V_{NMO}^2(0)\left(1+2\eta\right)}} , \end{aligned} \tag{A.10}$$

$$\nu_x = \frac{\mathrm{sgn}(\nu_x)V_{P0}\left(p_s\sqrt{1-s_\gamma^2}+p_r\sqrt{1-r_\gamma^2}\right)}{\sqrt{1-\frac{p_s^2V_{NMO}^2(0)}{1-2\eta p_s^2V_{NMO}^2(0)}}+\sqrt{1-\frac{p_r^2V_{NMO}^2(0)}{1-2\eta p_r^2V_{NMO}^2(0)}}} , \tag{A.11}$$

$$\nu_y = -\frac{V_{P0}\left(s_\gamma p_s + r_\gamma p_r\right)}{\sqrt{1-\frac{p_s^2V_{NMO}^2(0)}{1-2\eta p_s^2V_{NMO}^2(0)}}+\sqrt{1-\frac{p_r^2V_{NMO}^2(0)}{1-2\eta p_r^2V_{NMO}^2(0)}}} . \tag{A.12}$$

Note, that for the 2D problem we have $(s,r)_\gamma = 1$ and $-1 < (s,r)_\theta = V_{s,r}p_{s,r} < 1$, and that equations (A.10) and (A.12) then form a nonlinear system of two equations with two unknowns, viz. $p_{s,r}$.

Appendix B

Solving for scattering angle and azimuth for pre-stack demigration in homogeneous isotropic media

To find the angles $\theta_{s,r}$ (see Figure 2.1) and azimuths $\gamma_{s,r}$ (see Figure 2.3) for pre-stack demigration in homogeneous isotropic media, we set $\frac{dV}{d\theta} = 0$ and $V_s = V_r = V_{P0} = v$ in equations (2.61), (2.63)-(2.65). The resulting system of equations is

$$\sqrt{1-s_\gamma^2}\frac{s_\theta}{\sqrt{1-s_\theta^2}} = \sqrt{1-r_\gamma^2}\frac{r_\theta}{\sqrt{1-r_\theta^2}}\,, \tag{B.1}$$

$$\left(r_\gamma\frac{r_\theta}{\sqrt{1-r_\theta^2}} - s_\gamma\frac{s_\theta}{\sqrt{1-s_\theta^2}}\right) = \frac{4h}{vt_m}\,, \tag{B.2}$$

$$\mathrm{sgn}(\nu_x)\left(\frac{s_\theta\sqrt{1-s_\gamma^2}+r_\theta\sqrt{1-r_\gamma^2}}{\sqrt{1-s_\theta^2}+\sqrt{1-r_\theta^2}}\right) = \nu_x\,, \tag{B.3}$$

$$-\left(\frac{s_\theta s_\gamma + r_\theta r_\gamma}{\sqrt{1-s_\theta^2}+\sqrt{1-r_\theta^2}}\right) = \nu_y\,, \tag{B.4}$$

where $(s,r)_\theta = (\sin\theta_s, \sin\theta_r)$ and $(s,r)_\gamma = (\sin\gamma_s, \sin\gamma_r)$ are unknown. (The pathological cases $(s,r)_\theta = 1$, i.e. 90-degree dipping reflectors, are not included in this system of equations and we exclude these impractical cases by choosing $0 \le \theta < \pi/2$.)

To find s_θ, we first use equations (B.2) and (B.4) to eliminate $r_\theta r_\gamma$. Then we eliminate $s_\theta\sqrt{1-s_\gamma^2}$ from equations (B.1) and (B.3), and combine the results to give an equation

with only s_θ, s_γ and r_γ:

$$\left(\frac{\nu_x v t_m}{2}\right)^2 \left\{\nu_y \sqrt{1-s_\theta^2} + s_\theta s_\gamma\right\}^2 =$$
$$(1-r_\gamma^2)\left\{4h^2 + 2hvt_m\left(\nu_y + \frac{s_\theta s_\gamma}{\sqrt{1-s_\theta^2}}\right) + \left(\frac{s_\theta v t_m}{2}\right)^2\left(\nu_y + \frac{s_\theta s_\gamma}{\sqrt{1-s_\theta^2}}\right)^2\right\} . \tag{B.5}$$

In order to get r_γ^2 as a function of s_θ and s_γ, we use equations (B.1) and (B.2) to eliminate $r_\theta/\sqrt{1-r_\theta^2}$, and subsequently solve for r_γ^2. This gives

$$r_\gamma^2 = \frac{\left(2h\sqrt{1-s_\theta^2} + \frac{s_\theta s_\gamma v t_m}{2}\right)^2}{4h^2\left(1-s_\theta^2\right) + 2hs_\theta s_\gamma v t_m \sqrt{1-s_\theta^2} + \left(\frac{s_\theta v t_m}{2}\right)^2} . \tag{B.6}$$

Before we substitute this expression for r_γ^2 into equation (B.5), we first eliminate $r_\theta\sqrt{1-r_\gamma^2}$ from equations (B.1) and (B.3), and solve for s_γ^2, which gives

$$s_\gamma^2 = 1 + \nu_x^2 - \frac{\nu_x^2}{s_\theta^2} , \tag{B.7}$$

where $s_\theta \neq 0$. Then, using equations (B.6) and (B.7) in (B.5), gives

$$\left(1-s_\theta^2\right)\left\{h^2\left[\beta\left(1-s_\theta^2\right) - 4\nu_y^2\right] + 2\nu_y h\alpha\right\} =$$
$$\frac{v^2 t_m^2}{4}\left\{\beta\left[\left(1+2\nu_x^2\right)s_\theta^2 - \left(1+\nu_x^2\right)s_\theta^4 - \nu_x^2\right] - \nu_y^2\right\} , \tag{B.8}$$

with $\nu_x \neq 0$ and where we have defined

$$\alpha \equiv \frac{v t_m}{2}\left(1+\nu_x^2-\nu_y^2\right) , \tag{B.9}$$
$$\beta \equiv \left(1+\nu_x^2+\nu_y^2\right)^2 . \tag{B.10}$$

Note that the special case $\nu_x = 0$ is equivalent to the 2D case treated below. Equation

(B.8) is a quadratic equation in s_θ^2 that can be solved for s_θ to give

$$s_\theta = \sqrt{\frac{2h^2\left(\beta - 2\nu_y^2\right) + 2\nu_y h\left(\alpha \pm \gamma\right) + \left[\beta\left(\frac{vt_m}{2}\right)^2\left(1 + 2\nu_x^2\right) \mp \alpha\gamma\right]}{2\beta\left(h^2 + \left(1 + \nu_x^2\right)\left(\frac{vt_m}{2}\right)^2\right)}}, \tag{B.11}$$

with

$$\gamma \equiv \sqrt{4\nu_y^2 h^2 + \beta\left(\frac{vt_m}{2}\right)^2}. \tag{B.12}$$

The proper root in equation (B.11) is then found through substitution in the original system of equations (B.1)-(B.4).

Once s_θ is found, we need to solve for the remaining parameters r_θ and $(s,r)_\gamma$. To find r_θ we first use equation (B.4) in (B.2) to eliminate $r_\theta r_\gamma$ to give

$$-\frac{4h}{vt_m} = \nu_y + \frac{\nu_y\sqrt{1-s_\theta^2}}{\sqrt{1-r_\theta^2}} + s_\theta s_\gamma\left(\frac{1}{\sqrt{1-s_\theta^2}} + \frac{1}{\sqrt{1-r_\theta^2}}\right). \tag{B.13}$$

Then, to eliminate s_γ we use equation (B.7) and subsequently solve for r_θ to give

$$r_\theta = \sqrt{1 - \left(\frac{hvt_m\sqrt{s_\theta^2 - \nu_x^2(1-s_\theta^2)} - \sqrt{1-s_\theta^2}\left[\nu_y hvt_m + \Delta\right]}{4h^2 + 2\nu_y hvt_m + \Delta}\right)^2}, \tag{B.14}$$

where we have defined

$$\Delta \equiv \frac{v^2t_m^2}{4}\left(\nu_x^2 + \nu_y^2 - \frac{s_\theta^2}{1-s_\theta^2}\right).$$

Once $(s,r)_\theta$ are found, we solve equation (B.13) for s_γ to get

$$s_\gamma = -\frac{\sqrt{1-s_\theta^2}}{s_\theta}\frac{\left(\nu_y\sqrt{1-s_\theta^2} + \sqrt{1-r_\theta^2}\left(\frac{4h}{vt_m} + \nu_y\right)\right)}{\left(\sqrt{1-s_\theta^2} + \sqrt{1-r_\theta^2}\right)}, \tag{B.15}$$

provided $s_\theta \neq 0$. Finally, using equation (B.4) in (B.2) to eliminate $s_\gamma s_\theta$ we solve for r_γ.

$$r_\gamma = \frac{\sqrt{1-r_\theta^2}}{r_\theta}\frac{\left(\sqrt{1-s_\theta^2}\left(\frac{4h}{vt_m} - \nu_y\right) - \nu_y\sqrt{1-r_\theta^2}\right)}{\left(\sqrt{1-s_\theta^2} + \sqrt{1-r_\theta^2}\right)}, \tag{B.16}$$

provided $r_\theta \neq 0$.

B.0.1 Special cases

In 2D, the system of equations (B.1)-(B.4) reduces to two equations with two unknowns:

$$\left(\frac{r_\theta}{\sqrt{1-r_\theta^2}}-\frac{s_\theta}{\sqrt{1-s_\theta^2}}\right)=\frac{4h}{vt_m}\,, \tag{B.17}$$

$$-\left(\frac{s_\theta+r_\theta}{\sqrt{1-s_\theta^2}+\sqrt{1-r_\theta^2}}\right)=\nu_y\,, \tag{B.18}$$

with $-1<(s,r)_\theta<1$. In order to solve this system for the unknowns $(s,r)_\theta$, we first rewrite equation (B.17) to get

$$\sqrt{1-r_\theta^2}=\frac{r_\theta vt_m\sqrt{1-s_\theta^2}}{4h\sqrt{1-s_\theta^2}+s_\theta vt_m}\,. \tag{B.19}$$

Using this expression in equation (B.18) to eliminate $\sqrt{1-r_\theta^2}$ then gives

$$r_\theta=-\left(s_\theta+\frac{4\nu_y h\left(1-s_\theta^2\right)}{s_\theta vt_m+\sqrt{1-s_\theta^2}\left(4h+\nu_y vt_m\right)}\right)\,. \tag{B.20}$$

Then, using eq.(B.19) to calculate $\frac{1}{r_\theta^2}=\left(\sqrt{1-r_\theta^2}/r_\theta\right)^2+1$, and multiplying the result with r_θ^2 calculated using eq.(B.20), gives

$$\begin{aligned}&\left(\left[4h\sqrt{1-s_\theta^2}+s_\theta vt_m\right]^2+\left[1-s_\theta^2\right]v^2t_m^2\right)\\&\times\left(4h\nu_y\left[1-s_\theta^2\right]+s_\theta^2 vt_m+s_\theta\sqrt{1-s_\theta^2}\left[4h+\nu_y vt_m\right]\right)^2=\\&\left(4h\sqrt{1-s_\theta^2}+s_\theta vt_m\right)^2\left(s_\theta vt_m+\sqrt{1-s_\theta^2}\left[4h+\nu_y vt_m\right]\right)^2\end{aligned}\,. \tag{B.21}$$

Then dividing both sides of this expression by $\left(1-s_\theta^2\right)\left(4h\sqrt{1-s_\theta^2}+s_\theta vt_m\right)^2$ and simplifying the result, gives a quadratic equation in $\tan\theta_s=s_\theta/\sqrt{1-s_\theta^2}\equiv\tau_{\theta s}$, viz.

$$\nu_y vt_m\tau_{\theta s}^2+\left(4h\nu_y+vt_m\left(\nu_y^2-1\right)\right)\tau_{\theta s}+2h\left(\nu_y^2-1\right)-\nu_y vt_m=0\,, \tag{B.22}$$

with roots

$$\tau_{\theta s} = \frac{1}{2\nu_y}\left(1-\nu_y^2\right) - \frac{2h}{vt_m} \pm \sqrt{\frac{4h^2}{v^2t_m^2} + \frac{\left(1+\nu_y^2\right)^2}{4\nu_y^2}}\,. \tag{B.23}$$

Then, by using this expression in equation (B.17), we find

$$\tau_{\theta r} = \frac{1}{2\nu_y}\left(1-\nu_y^2\right) + \frac{2h}{vt_m} \pm \sqrt{\frac{4h^2}{v^2t_m^2} + \frac{\left(1+\nu_y^2\right)^2}{4\nu_y^2}}\,. \tag{B.24}$$

Therefore, $(s,r)_\theta$ are then given by

$$(s,r)_\theta = \sin\left(\arctan\left(\tau_{\theta s},\tau_{\theta r}\right)\right)\,. \tag{B.25}$$

The proper roots in equation (B.23) and (B.24) are then chosen through substitution in the original system of equations (B.17) and (B.18). Note that the pathological cases $(s,r)_\theta = 0$ mentioned in the previous subsection are included in this solution.

For the special case $\nu_y = 0$, i.e., the 2D zero-dip case, the solution for $(s,r)_\theta$ is given simply by

$$s_\theta = \sin\left(\arctan\left(\frac{-2h}{vt_m}\right)\right) = -r_\theta\,. \tag{B.26}$$

Appendix C

Explicit expressions for map DMO in isotropic homogeneous media

Map dip moveout (DMO) is the mapping from the midpoint location y_u, two-way traveltime t_u, and slope p_u in a common-offset section, given a velocity, to their counterparts after DMO, i.e., the midpoint location y_0, two-way traveltime τ_0, and slope p_m^{DMO}. In order to derive the map DMO equations for an isotropic homogeneous medium, we start with the DMO ellipse in such a medium, just as we started with the pre-stack migration ellipse when we derived the pre-stack map migration equations in Chapter 2. The DMO ellipse is given by

$$\frac{(y_0 - y_u)^2}{h^2} + \frac{\tau_0^2}{t_u^2 - \frac{4h^2}{v^2}} = 1\,, \tag{C.1}$$

where y_0 and τ_0 are the midpoint location and two-way traveltime, respectively, after DMO, y_u and t_u are the midpoint location and two-way traveltime before DMO, h is the half-offset, and v is the medium velocity (assumed constant). Calculating the derivative of equation (C.1) with respect to y_0 (i.e. $\frac{d}{dy_0}$) and subsequently solving for $p_m^{DMO} := \frac{1}{2}\frac{d\tau_0}{dy_0}$, we get

$$p_m^{DMO} = \frac{(y_u - y_0)\left(t_u^2 - \frac{4h^2}{v^2}\right)}{2h^2\tau_0}\,. \tag{C.2}$$

Similarly, calculating the derivative of the DMO ellipse with respect to y_u we get

$$\frac{y_u - y_0}{h^2} - \frac{2\tau_0^2 t_u p_u}{\left(t_u - \frac{4h^2}{v^2}\right)^2} = 0\,, \tag{C.3}$$

where $p_u := \frac{1}{2}\frac{dt_u}{dy_u}$. Equations (C.1), (C.2), and (C.3) form a system of three equations with three unknowns (assuming the velocity is known) for y_0, τ_0, and p_m^{DMO}. Hence, the expressions for map DMO follow from solving equations (C.1), (C.2), and (C.3) for y_0, τ_0, and p_m^{DMO}.

Rewriting equation (C.3) gives

$$y_u - y_0 = \frac{2h^2\tau_0^2 t_u p_u}{\left(t_u^2 - \frac{4h^2}{v^2}\right)^2}. \tag{C.4}$$

Eliminating the term $y_u - y_0$ in (C.1) using equation (C.4) and subsequently solving for τ_0, then gives

$$\tau_0 = \frac{\left(t_u^2 - \frac{4h^2}{v^2}\right)^{3/2}}{2\sqrt{2}ht_u p_u} \sqrt{\sqrt{1 + \frac{16h^2 t_u^2 p_u^2}{\left(t_u^2 - \frac{4h^2}{v^2}\right)^2}} - 1}\,, \tag{C.5}$$

where we ignored the other root for τ_0 as it is complex. Then, using equation (C.5) in (C.4) and solving for y_0 gives the explicit expression for y_0, viz.

$$y_0 = y_u - \frac{\left(t_u^2 - \frac{4h^2}{v^2}\right)}{4t_u p_u} \left[1 - \sqrt{1 + \frac{16h^2 t_u^2 p_u^2}{\left(t_u^2 - \frac{4h^2}{v^2}\right)^2}}\right]. \tag{C.6}$$

Note that if we set $h = 0$, we indeed get $y_0 = y_u$ as expected, since in this case applying DMO to the data leaves the data unchanged. Finally, using equations (C.5) and (C.6) in equation (C.2) gives

$$p_m^{DMO} = -\frac{\sqrt{t_u^2 - \frac{4h^2}{v^2}}}{2\sqrt{2}h} \sqrt{1 - \sqrt{1 + \frac{16h^2 t_u^2 p_u^2}{\left(t_u^2 - \frac{4h^2}{v^2}\right)^2}}}\,. \tag{C.7}$$

Equations (C.5), (C.6), and (C.7) are the 2D map DMO equations for an isotropic homogeneous medium.

Appendix D

Explicit expressions for inverse map DMO in isotropic homogeneous media

Inverse map DMO is the mapping from the midpoint location y_0, two-way traveltime τ_0, and slope p_m^{DMO} after DMO, to their counterparts before DMO, i.e., the midpoint location y_u, two-way traveltime t_u, and slope p_u. To determine the explicit expressions of y_u, t_u, and p_u as functions of y_0, τ_0, and p_m^{DMO}, we use the same system of equations as we used to derive the map DMO equations in Appendix C, i.e., equations (C.1), (C.2), and (C.3). We now treat y_u, t_u, and p_u as unknowns and solve for the unmigrated specular reflection (y_u, t_u, p_u) as a function of y_0, τ_0, and p_m^{DMO}, given a velocity v and half-offset h.

Rewriting equation (C.2) we get

$$y_u - y_0 = \frac{2h^2 \tau_0 p_m^{DMO}}{t_n^2}\,, \tag{D.1}$$

where we have defined $t_n^2 := t_u^2 - \dfrac{4h^2}{v^2}$. Using equation (D.1) in equation (C.1) and solving for t_u, we get

$$t_u = \sqrt{\frac{4h^2}{v^2} + \frac{\tau_0^2}{2}\left(1 + A\right)}\,, \tag{D.2}$$

where we have defined $A := A(p_m^{DMO}, \tau_0, h) = \sqrt{1 + \dfrac{16h^2\left(p_m^{DMO}\right)^2}{\tau_0^2}}$, and we chose the proper root using the fact that for $h = 0$ we must have $t_u = \tau_0$ (since in this case DMO, and its inverse, leave the data unchanged). Using equation (D.2) in (C.1) and solving for y_u, we find

$$y_u = y_0 - h \operatorname{sgn} p_m^{DMO} \frac{\sqrt{A-1}}{\sqrt{A+1}}\,. \tag{D.3}$$

Then, solving equation (C.3) for p_u we find

$$p_u = \left(\frac{y_u - y_0}{h^2}\right) \frac{\left(t_u^2 - \dfrac{4h^2}{v^2}\right)^2}{2\tau_0^2 t_u}\,. \tag{D.4}$$

Finally, we find the explicit expression for p_u in terms of y_0, τ_0, and p_m^{DMO}, by using

equations (D.2) and (D.3) in (D.4). This gives

$$p_u = \operatorname{sgn} p_m^{DMO} \frac{\tau_0^2}{2h} \frac{\sqrt{A-1}\,(A+1)^{3/2}}{\sqrt{\dfrac{4h^2}{v^2} + \dfrac{\tau_0^2}{2}(A+1)}}, \tag{D.5}$$

where we used that $\operatorname{sgn} p_u = \operatorname{sgn} p_m^{DMO}$. Equations (D.2), (D.3) and (D.5) define the inverse map DMO equations in an isotropic homogeneous medium.

Appendix E

Comparison with Cohen's (1998) expressions for 2D zero-offset map time-migration

Cohen (1998) presented equations for the position of a (straight) ray in a vertical symmetry plane of a homogeneous anisotropic medium. In this appendix, I show that these equations are equivalent to the zero-offset map time-migration equations for the migrated location, assuming propagation indeed occurs in a vertical symmetry plane of a homogeneous anisotropic medium. Cohen's expressions [his equation (28)] for the position along a straight ray are given by

$$x = g_1 \tau \,, \quad z = g_3 \tau \tag{E.1}$$

where x and z are the horizontal and vertical position along the ray, τ is the one-way traveltime, and g_1 and g_3 are the horizontal and vertical components, respectively, of the group-velocity vector, given by [Cohen's equations (26) and (27)],

$$g_1 = -\frac{dq}{dp} g_3 \,, \quad g_3 = \frac{1}{q - p\dfrac{dq}{dp}} \,. \tag{E.2}$$

In these expressions, p and q are the horizontal and vertical components of the slowness-vector, respectively. From equation (E.1) we can deduce, by substituting $z := z_m = V_{p0} t_m/2$ and $\tau := t_u/2$, that

$$t_m = \frac{t_u}{V_{p0}\left(q - p\dfrac{dq}{dp}\right)} \,, \tag{E.3}$$

where t_m is the migrated two-way traveltime, and t_u is the unmigrated two-way traveltime. This equation is valid in any vertical symmetry plane of a homogeneous medium. Although Cohen does not explicitly call this equation an equation for map time-migration, it indeed is since it makes explicit use of the slope information in the data to find the migrated position in time.

Since the vertical slowness can be written as $q = \sqrt{1/V(\theta)^2 - p^2}$ with $V(\theta)$ the phase velocity, we have

$$\frac{dq}{dp} = -\frac{V(\theta)}{\sqrt{1 - p^2 V(\theta)^2}} \left(\frac{1}{V(\theta)^3} \frac{dV}{dp} + p \right) \,. \tag{E.4}$$

To evaluate this expression, we need to find the derivative of the phase-velocity with respect

to the horizontal slowness component p. Using the chain rule we have

$$\frac{dV}{dp} = \frac{dV}{d\theta}\frac{d\theta}{dp} = \frac{dV}{d\theta}\left(\frac{dp}{d\theta}\right)^{-1} . \tag{E.5}$$

Since in general $p = \sin\theta/V$, we find that

$$\frac{dV}{dp} = \frac{dV}{d\theta}\frac{V(\theta)}{\sqrt{1-p^2V(\theta)^2} - p\dfrac{dV}{d\theta}} . \tag{E.6}$$

Using this expression in equation (E.4) and substituting the result in equation (E.3) then gives

$$t_m = t_u\frac{V(\theta)}{V_{p0}}\left(\sqrt{1-p^2V(\theta)^2} - p\frac{dV}{d\theta}\right) , \tag{E.7}$$

which is identical to equation (3.1) in the main text, when we substitute p_u for p.

From equations (E.1) and (E.2) we deduce (upon substituting $x := x_u - x_m$ and $\tau := t_u/2$) that

$$x_u - x_m = \frac{\dfrac{t_u}{2}\dfrac{dq}{dp}}{p\dfrac{dq}{dp} - q} . \tag{E.8}$$

Using equations (E.4) and (E.6) in equation (E.8), we find

$$x_m = x_u - \frac{V^2(\theta)p\,t_u}{2} - \frac{V(\theta)p\,t_u}{2}\frac{dV}{d\theta}\sqrt{\frac{1}{V^2(\theta)p^2} - 1} , \tag{E.9}$$

which is identical to equation (3.2) in the main text (provided we again substitute p_u for p).

Appendix F

Algebraic solution to the general cubic equation

The algebraic solution to the general cubic equation was first published in the *Ars Magna* by Cardano (1545), although he obtained the solution most likely from Tartaglia and del Ferro (Boyer, 1991, p.282-283). The method for solving the cubic presented here closely follows the treatment by Richardson (1958, p.327-329).

A general cubic equation is given by

$$x^3 + a_2 x^2 + a_1 x^1 + a_0 = 0\,. \tag{F.1}$$

To solve equation (F.1), we first substitute

$$x := z - \lambda\,, \tag{F.2}$$

and solve for λ such that the coefficient in front of the z^2 term is zero. To achieve this, λ must satisfy $-3\lambda + a_2 = 0$, giving

$$\lambda = \frac{a_2}{3}\,. \tag{F.3}$$

Using the definition of z from equation (F.2) together with the value of λ just determined, the cubic equation (F.1) becomes

$$z^3 + pz + q = 0\,, \tag{F.4}$$

where we defined

$$p := a_1 - \frac{a_2^2}{3}\,, \tag{F.5}$$

$$q := a_0 - \frac{a_1 a_2}{3} + \frac{2}{27}a_2^3\,. \tag{F.6}$$

Equation (F.4) is usually referred to as the *depressed* cubic.

To solve the depressed cubic equation (F.4), we use Viéta's "magic" substitution

$$z := u - \frac{p}{3u}\,, \quad u \neq 0\,. \tag{F.7}$$

Using this substitution in equation (F.4), we find

$$u^6 + qu^3 - \frac{p^3}{27} = 0\,, \tag{F.8}$$

which is a quadratic equation in u^3 and has solutions

$$u^3 = -\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}\,, \tag{F.9}$$

and

$$u^3 = -\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}\,. \tag{F.10}$$

Let α be a solution to u in equation (F.9) and β a solution to u in equation (F.10). Then, defining $\omega := -\frac{1}{2} + \frac{\sqrt{3}}{2}i$, we know that the six combined solutions to equations (F.9) and (F.10) are $\alpha, \omega\alpha, \omega^2\alpha, \beta, \omega\beta$, and $\omega^2\beta$. Also, we have

$$\left(-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}\right)\left(-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}\right) = -\frac{p^3}{27}\,, \tag{F.11}$$

which implies that the solutions α and β to equations (F.9) and (F.10), must satisfy

$$\alpha\beta = -\frac{p}{3}\,. \tag{F.12}$$

Hence, using the three cube roots $\alpha, \omega\alpha$, and $\omega^2\alpha$ of equation (F.9) in definition (F.7), the three roots to the depressed cubic equation (F.4) are given by

$$z_1 = \alpha - \frac{p}{3\alpha} = \alpha - \frac{p}{3\frac{-p}{3\beta}} = \alpha + \beta\,, \tag{F.13}$$

$$z_2 = \omega\alpha - \frac{p}{3\omega\alpha} = \omega\alpha - \frac{1}{\omega}\frac{p}{3\frac{-p}{3\beta}} = \omega\alpha + \omega^2\beta\,, \tag{F.14}$$

$$z_3 = \omega^2\alpha - \frac{p}{3\omega^2\alpha} = \omega^2\alpha - \frac{1}{\omega^2}\frac{p}{3\frac{-p}{3\beta}} = \omega^2\alpha + \omega\beta\,, \tag{F.15}$$

where we used equation (F.12) and $\omega^3 = 1$. Note that if we had chosen to use the three cube roots $\beta, \omega\beta$, and $\omega^2\beta$ in definition (F.7), we would have found the same three roots z_1, z_2, and z_3. This means that the sign of the square root in equations (F.9) and (F.10) is irrelevant to the final roots of the general cubic equation. If $\alpha = 0$ (which happens when $p = 0$, $q \neq 0$, and $q > 0$), we simply choose the three cube roots $\beta, \omega\beta$, and $\omega^2\beta$ in definition (F.7).

The roots x_1, x_2, and x_3 to the general cubic equation are then found using equations (F.13) - (F.15) and (F.3) in equation (F.2). This gives

$$x_1 = \alpha + \beta - \frac{a_2}{3}\,, \quad x_2 = \omega\alpha + \omega^2\beta - \frac{a_2}{3}\,, \quad x_3 = \omega^2\alpha + \omega\beta - \frac{a_2}{3}\,. \tag{F.16}$$

When $p = 0$ and $q = 0$, the depressed cubic equation is simply $z^3 = 0$, giving the roots to the general cubic equation $x_1 = x_2 = x_3 = -\frac{a_2}{3}$.

Appendix G

Algebraic solution to the general quartic equation

The solution to the general quartic equation is originally due to Ferrari, but was first published in the *Ars Magna* by Cardano (1545). In this appendix we follow the treatment by Richardson (1958, p.329-331), although we add the explicit equations for the roots of the quartic equation in terms of the coefficients a_3, a_2, a_1, and a_0.

The general quartic equation is given by

$$x^4 + a_3x^3 + a_2x^2 + a_1x + a_0 = 0\,. \tag{G.1}$$

To solve the quartic equation, we aim to write the equation in the form $P^2 = Q^2$, where both P and Q are quadratic functions of x. In this way the quartic equation is reduced to two quadratic equations (i.e. $P = \pm Q$), and the four solutions to these two quadratic equations are the solutions to the general quartic equation. The quartic equation is therefore sometimes also referred to as a *biquadratic* equation.

Rewriting equation (G.1) and adding a term $(a_3x/2)^2$ on both sides, gives

$$\left(x^2 + \frac{a_3x}{2}\right)^2 = \left(\frac{a_3^2}{4} - a_2\right)x^2 - a_1x - a_0\,. \tag{G.2}$$

The right-hand side of equation (G.2) is generally not a perfect square. Hence, in order for both the left- and right-hand sides of equation (G.2) to be a perfect square, we add the terms $\left(x^2 + \frac{a_3x}{2}\right)y + \frac{y^2}{4}$ to both sides. Doing this gives

$$\left(x^2 + \frac{a_3x}{2} + \frac{y}{2}\right)^2 = \left(\frac{a_3^2}{4} - a_2 + y\right)x^2 + \left(\frac{a_3y}{2} - a_1\right)x + \frac{y^2}{4} - a_0\,. \tag{G.3}$$

The right-hand side of this equation is a perfect square if its determinant $D = \left(\frac{a_3y}{2} - a_1\right)^2 - 4\left(\frac{a_3^2}{4} - a_2 + y\right)\left(\frac{y^2}{4} - a_0\right)$ equals zero. Setting $D = 0$ leads to the resolvent cubic equation in y, given by

$$y^3 - a_2y^2 + (a_1a_3 - 4a_0)\,y + 4a_0a_2 - a_1^2 - a_0a_3^2 = 0\,. \tag{G.4}$$

This equation is called the resolvent cubic equation, because any of its roots turn the right-hand side of equation (G.3) into a perfect square. Such a root can be found using the method outlined in appendix F for solving a cubic equation.

Denoting y_0 as a root to the resolvent cubic equation (G.4) and defining $R := \sqrt{\frac{a_3^2}{4} - a_2 + y_0}$, we need to separate two independent cases: I) $R \neq 0$, and II) $R = 0$. In the first case (I), we rewrite equation (G.3) as

$$\left(x^2 + \frac{a_3 x}{2} + y2\right)^2 = \left(Rx + \frac{a_3 y_0 - 2a_1}{4R}\right)^2 , \tag{G.5}$$

which can be simplified to the two quadratic equations

$$x^2 + \left(\frac{a_3}{2} \pm R\right) x + y_0 2 \pm \frac{a_3 y_0 - 2a_1}{4R} = 0 . \tag{G.6}$$

In this case (i.e. $R \neq 0$), the explicit expressions for the solutions of the general quartic equation (G.1) are given by

$$x_1 = -\frac{a_3}{4} - \frac{1}{2}\sqrt{\frac{a_3^2}{4} - a_2 + y_0} + \ldots \\ \frac{1}{2}\sqrt{\left(\frac{a_3}{2} + \sqrt{\frac{a_3^2}{4} - a_2 + y_0}\right)^2 - 2y_0 - \frac{(a_3 y_0 - 2a_1)}{\sqrt{\frac{a_3^2}{4} - a_2 + y_0}}} , \tag{G.7}$$

$$x_2 = -\frac{a_3}{4} - \frac{1}{2}\sqrt{\frac{a_3^2}{4} - a_2 + y_0} - \ldots \\ \frac{1}{2}\sqrt{\left(\frac{a_3}{2} + \sqrt{\frac{a_3^2}{4} - a_2 + y_0}\right)^2 - 2y_0 - \frac{(a_3 y_0 - 2a_1)}{\sqrt{\frac{a_3^2}{4} - a_2 + y_0}}} , \tag{G.8}$$

$$x_3 = -\frac{a_3}{4} + \frac{1}{2}\sqrt{\frac{a_3^2}{4} - a_2 + y_0} + \ldots \\ \frac{1}{2}\sqrt{\left(\frac{a_3}{2} - \sqrt{\frac{a_3^2}{4} - a_2 + y_0}\right)^2 - 2y_0 + \frac{(a_3 y_0 - 2a_1)}{\sqrt{\frac{a_3^2}{4} - a_2 + y_0}}} , \tag{G.9}$$

$$x_4 = -\frac{a_3}{4} + \frac{1}{2}\sqrt{\frac{a_3^2}{4} - a_2 + y_0} - \ldots \\ \frac{1}{2}\sqrt{\left(\frac{a_3}{2} - \sqrt{\frac{a_3^2}{4} - a_2 + y_0}\right)^2 - 2y_0 + \frac{(a_3 y_0 - 2a_1)}{\sqrt{\frac{a_3^2}{4} - a_2 + y_0}}} . \tag{G.10}$$

In case $R = 0$ (case II), we rewrite equation (G.3) as

$$\left(x^2 + \frac{a_3 x}{2} + y2\right)^2 = \left(\sqrt{\frac{y_0^2}{4} - a_0}\right)^2 , \tag{G.11}$$

which can be simplified to the two quadratic equations

$$x^2 + \frac{a_3}{2}x + \frac{y_0}{2} \pm \sqrt{\frac{y_0^2}{4} - a_0} = 0 . \tag{G.12}$$

In this case (i.e. $R = 0$), the explicit expressions for the solutions of the general quartic equation (G.1) are given by

$$x_1 = -\frac{a_3}{4} + \frac{1}{2}\sqrt{\frac{a_3^2}{4} - 2y_0 - 4\sqrt{\frac{y_0^2}{4} - a_0}} , \tag{G.13}$$

$$x_2 = -\frac{a_3}{4} - \frac{1}{2}\sqrt{\frac{a_3^2}{4} - 2y_0 - 4\sqrt{\frac{y_0^2}{4} - a_0}} , \tag{G.14}$$

$$x_3 = -\frac{a_3}{4} + \frac{1}{2}\sqrt{\frac{a_3^2}{4} - 2y_0 + 4\sqrt{\frac{y_0^2}{4} - a_0}} , \tag{G.15}$$

$$x_4 = -\frac{a_3}{4} - \frac{1}{2}\sqrt{\frac{a_3^2}{4} - 2y_0 + 4\sqrt{\frac{y_0^2}{4} - a_0}} , \tag{G.16}$$

Appendix H

A particular construction of a tight frame of curvelets

In this appendix, we give a detailed explanation of the construction of curvelets and show that they constitute a tight-frame for functions in $L^2(\mathbb{R}^2)$. This treatment closely follows the construction of real-valued curvelets by Candès & Donoho (2004b) except for some added derivations and explanations to clarify the construction of curvelets for the non-specialist. We deviate in some places from the original treatment to clarify certain details. For example, we immediately use the notion of *splitting at every other scale* (explained later). A frame similar to the tight frame of curvelets was earlier introduced by Smith (1998).

Before treating the construction of curvelets, we mention that we want to be able to reconstruct a signal in a function space $\mathcal{X}$; i.e., we want curvelets to satisfy the reconstruction formula

$$f = \sum_{\mu \in M} (f, c_\mu)\, c_\mu \quad , \tag{H.1}$$

where c_μ denotes a curvelet with multi-index μ (and M some index-set), and $f \in \mathcal{X}$. It is known [e.g., Hernández & Weiss (1996, pp.334-336) and Appendix I in this paper] that if $\mathcal{X}$ is a Hilbert space and c_μ are elements in this space, equation (H.1) is satisfied if and only if

$$\|f\|_{\mathcal{X}}^2 = \sum_{\mu \in M} |(f, c_\mu)|^2 \ , \quad \forall f \in \mathcal{X} \ , \tag{H.2}$$

where $\|\cdot\|_{\mathcal{X}}$ denotes the norm on $\mathcal{X}$. The latter expression implies that the energy of the signal f is conserved through the decomposition (H.1); in other words, the curvelets c_μ should be a *partition of unity* (i.e., unit energy) for the reconstruction formula (H.1) to hold. Condition (H.2) is the definition of a normalized tight frame (see Appendix I for an explanation of tight frames).

Curvelets can be thought of as 2D (anisotropic) extensions to wavelets. Just as in 1D wavelets are localized in one variable and its Fourier dual, curvelets are localized in two variables *and* their two Fourier duals. Such localization is understood within the limits imposed by the Heisenberg uncertainty principle, i.e., $\Delta t \Delta f \geq 1$, with t and f denoting time and frequency, respectively. This principle thus determines a *tile* in the space with t along one axis and f along the other. For curvelets in 2D, choosing the variables to be x and z, such localization is obtained through (hyperdimensional) tiling of the spatial domain

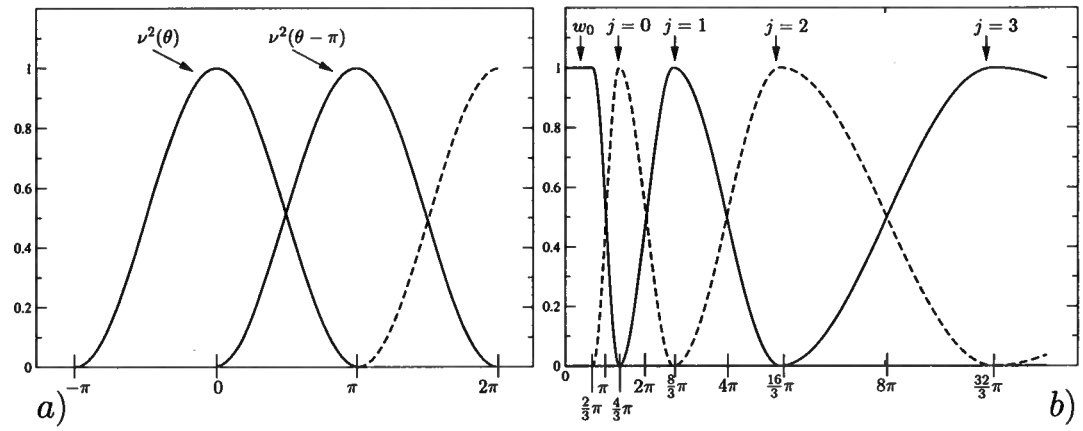


Figure H.1. Example of a possible angular window function $\nu(\theta)$ (a), and Lemarié-Meyer windows $|w(2^{-j}|\boldsymbol{k}|)|^2$ for the radial windows (b). The Lemarié-Meyer windows for j and $j+1$ are equal at $2^{j+1}\pi$.

(x,z) and the spectral domain (k_x,k_z). We treat the spectral localization first followed by the spatial one.

Let $\chi_{j,l}(\boldsymbol{k})$ be a window (or tile) in the spectral domain, with j an index related to the radial (i.e., scale) direction, and l an index related to the angular direction; i.e. the localization in the spectrum is dealt with in polar coordinates r and θ. In order for curvelets to constitute a tight frame and thus satisfy the reconstruction formula (H.1) (see Appendix I), the windows $\chi_{j,l}(\boldsymbol{k})$ must satisfy

$$\sum_j \sum_l |\chi_{j,l}(\boldsymbol{k})|^2 = 1,\ j \in J, l \in L, \tag{H.3}$$

such that $\chi_{j,l}(\boldsymbol{k})$ indeed is a partition of unity. The window $\chi_{j,l}(\boldsymbol{k})$ is constructed through multiplication of an angular window $\nu(\theta)$ and a radial window $w(|\boldsymbol{k}|)$, where $\boldsymbol{k}$ is the wave-vector. We treat the angular window first, followed by the radial window.

Let $\nu(\theta)$ be an even, real-valued, non-negative, angular window function that is infinitely continuously differentiable (i.e., C^∞) and compactly supported on $[-\pi,\pi)$. This function is chosen to be 2π-periodic, such that $\nu(\theta)$ and $\nu(\theta+\pi)$ are exactly out of phase; Figure H.1a shows a possible choice of such a window function. Then the angular window satisfies

$$|\nu(\theta)|^2 + |\nu(\theta+\pi)|^2 = 1,\ \theta \in [0,2\pi).$$

Defining the windows $\nu_{j,l}(\theta) = \nu(2^j\theta - l\pi)$, with $j \geq 0$ and $l = 0,1,2,\ldots,2^j-1$, it follows that these windows cover the interval $[-\pi/2^j,\pi)$. The windows $\nu_{j,l}(\theta+\pi)$ then cover $[-\pi-\pi/2^j,0)$. Taking into account the 2π-periodicity of $\nu(\theta)$, it follows that $\nu_{j,l}(\theta)$ and

$\nu_{j,l}(\theta+\pi)$ together cover $[0, 2\pi)$. It is important to note that each window $\nu_{j,l}(\theta)$ has width $2\pi/2^j$ and that the width of the window thus depends on the scale j.

Due to the 2π-periodicity of $\nu(\theta)$, we have

$$\nu_{j,l}(\theta+\pi) = \nu(2^j(\theta+\pi) - l\pi) = \nu\left(2^j\theta - \left(l+2^j\right)\pi\right) = \nu_{j,l+2^j}(\theta). \tag{H.4}$$

Using this, it follows that for each $j \geq 0$

$$\begin{aligned}
\sum_{l=0}^{2^{j+1}-1} |\nu_{j,l}(\theta)|^2 &= \sum_{l=0}^{2^j-1} |\nu_{j,l}(\theta)|^2 + \sum_{l=2^j}^{2^{j+1}-1} |\nu_{j,l}(\theta)|^2 \\
&= \sum_{l=0}^{2^j-1} |\nu_{j,l}(\theta)|^2 + \sum_{l=0}^{2^j-1} \left|\nu_{j,l+2^j}(\theta)\right|^2 \\
&= \sum_{l=0}^{2^j-1} \left(|\nu_{j,l}(\theta)|^2 + |\nu_{j,l}(\theta+\pi)|^2\right) = 1\,.
\end{aligned} \tag{H.5}$$

Therefore, energy is indeed conserved in the angular direction using the window functions $\nu_{j,l}(\theta)$. Another way of saying this is that the windows $\nu_{j,l}(\theta)$ satisfy the *admissibility condition* (Candès *et al.*, 2005). There are many different window functions that satisfy this admissibility condition. Hence, curvelets can be constructed using different choices of angular window functions, much as in the construction of wavelets.

The angular windowing provides the curvelets with geometry, i.e., with their main associated directions. To give curvelets their bandlimited nature, howver, the (radial) frequency axis is subdivided into bandpass filters (i.e., subbands). This is done using a radial window w. A possible choice for such a window is the Lemarié-Meyer window [see Hernández & Weiss (1996, p.27-28) and Figure H.1b]. Let w be a C^∞ real-valued, non-negative, function with support included in $[2\pi/3, 8\pi/3]$. Then Meyer introduces a partition of unity as

$$|w_0(r)|^2 + \sum_{j\geq 0} \left|w(2^{-j}r)\right|^2 = 1, \quad \forall r \geq 0, \tag{H.6}$$

with w_0 a *coarse scale* C^∞ window that equals 1 on $[0, 2\pi/3]$ and vanishes beyond $4\pi/3$.

Let $\boldsymbol{k}$ denote the (angular) frequency vector, i.e., $\boldsymbol{k} = (k_x, k_z)$. Then, combining the angular and radial windows into one window $\chi_{j,l}(\boldsymbol{k})$ as

$$\chi_{j,l}(\boldsymbol{k}) = w\left(2^{-j}\,|\boldsymbol{k}|\right)\left[\nu_{\lfloor j/2\rfloor,l}(\theta) + \nu_{\lfloor j/2\rfloor,l}(\theta+\pi)\right], \tag{H.7}$$

and setting $|\chi_0(\boldsymbol{k})|^2 = |w_0(|\boldsymbol{k}|)|^2 + |w(|\boldsymbol{k}|)|^2$, it follows that

$$
\begin{aligned}
&|\chi_0(\boldsymbol{k})|^2 + \sum_{l=0,j\geq 1}^{2^{\lfloor j/2\rfloor}-1} |\chi_{j,l}(\boldsymbol{k})|^2 \\
&\quad = |w_0(|\boldsymbol{k}|)|^2 + |w(|\boldsymbol{k}|)|^2 + \sum_{l=0,j\geq 1}^{2^{\lfloor j/2\rfloor}-1} \left|w\left(2^{-j}\,|\boldsymbol{k}|\right)\right|^2 \left|\nu_{\lfloor j/2\rfloor,l}(\theta) + \nu_{\lfloor j/2\rfloor,l}(\theta+\pi)\right|^2 \\
&\quad = |w_0(|\boldsymbol{k}|)|^2 + |w(|\boldsymbol{k}|)|^2 + \sum_{j\geq 1} \left|w\left(2^{-j}\,|\boldsymbol{k}|\right)\right|^2 \sum_{l=0}^{2^{\lfloor j/2\rfloor}-1} \left[\nu_{\lfloor j/2\rfloor,l}(\theta) + \nu_{\lfloor j/2\rfloor,l}(\theta+\pi)\right] \\
&\quad = |w_0(|\boldsymbol{k}|)|^2 + \sum_{j\geq 0} \left|w\left(2^{-j}\,|\boldsymbol{k}|\right)\right|^2 = 1\,, \qquad \text{(H.8)}
\end{aligned}
$$

where we used equations (H.5) and (H.6). Here, the notation $\lfloor j \rfloor$ denotes the integer part of j. From equation (H.8), $\chi_{j,l}(\boldsymbol{k})$ [together with $\chi_0(\boldsymbol{k})$] form a partition of unity of the whole frequency plane. Note that by construction, the $\chi_{j,l}(\boldsymbol{k})$ have compact support, i.e., curvelets are compactly supported in the spectral domain. Note that there are many different radial windows that could form such a partition of unity, just as for the angular windows; i.e., many different variants of curvelets can be constructed by using different radial window functions.

From the definition of $\chi_{j,l}(\boldsymbol{k})$, $\chi_{j,l}(\boldsymbol{k})$ is obtained through a rotation of $\chi_{j,0}(\boldsymbol{k})$. With this observation, the tiling $\chi_{j,l}(\boldsymbol{k})$ of the frequency plane (k_x, k_z) can be depicted as seen in Figure 4.1a in Chapter 4. Note that the definition of $\chi_{j,l}(\boldsymbol{k})$ using $\nu_{\lfloor j/2\rfloor,l}$ in equation (H.7), means that the number of angular windows increases only every other scale (see Figure 4.1a). This *splitting at every other scale* is a consequence of the choice of dyadic subbands in combination with the parabolic scaling. With the subband filtering being dyadic (i.e., the length of the support of the radial window being multiplied by 2 each time the scale index increases by 1), the width of the curvelet is multiplied by $\sqrt{2}$ each time the scale index increases by 1. Hence, the only way to have an integer number of angular windows per subband with the windows satisfying the parabolic sclaling, is to double the number of angular windows at every other scale only.

Since the radial (or subband) windows $w(2^{-j}|\boldsymbol{k}|)$ for j and $j+1$ are equal at $2^{j+1}\pi$ (see Figure H.1b), the frequencies of $\chi_{j,l}(\boldsymbol{k})$ are supported near the *dyadic coronae* $\{2^j\pi \leq |\boldsymbol{k}| \leq 2^{j+1}\pi\}$, depicted in Figure 4.1a. These dyadic coronae are directly related to the familiar dyadic subband filtering in wavelet theory. This decomposition of the frequency plane is the same as the second dyadic decomposition treated by Stein (1993) in relation to oscillatory integrals of the second kind (or Fourier Integral Operators).

Ignoring the splitting at every other scale, we see that the length of an angular wedge (i.e., in the radial direction) is proportional to $2^{j+1} - 2^j = 2^j$, while the width of the wedge is proportional to $2^j \times 2^{-j/2} = 2^{j/2}$. This means that the volume of the angular wedge is $O\left(2^j \times 2^{j/2}\right)$. Considering the Heisenberg uncertainty principle, a curvelet is therefore supported in the spatial domain near a rectangle of width proportional to 2^{-j} and length

$2^{-j/2}$; i.e., the width of this rectangle is the inverse of the length of the angular wedge, while the length of the rectangle is the inverse of the width of the wedge. Therefore, curvelets obey the anisotropic scaling relation

$$\text{width} \propto \text{length}^2 \,, \tag{H.9}$$

where proportionality is used to indicate the omitted constants[1]. This scaling is referred to as the *parabolic scaling*. Note that the scaling relation (H.9) holds in the spatial domain, while in the frequency domain we have length $\propto$ width^2.

So far we have shown the tiling of the frequency plane with curvelets. The windows $\chi_{j,l}(\boldsymbol{k})$ allow us to rotate and dilate curvelets. To be able to translate curvelets in the spatial domain, the windows $\chi_{j,l}(\boldsymbol{k})$ are multiplied with a local (i.e., over the spectral support of a curvelet) L^2 orthonormal basis. Since all windows $\chi_{j,l}(\boldsymbol{k})$ are simply rotated versions of $\chi_{j,0}(\boldsymbol{k})$, we first consider the window $\chi_{j,0}(\boldsymbol{k})$. The support of this window is contained in the rectangle $R = I_{1j} \times I_{2j}$, with

$$I_{1j} = \{k_x, t_j \leq k_x \leq t_j + L_j\}, \quad I_{2j} = \{k_z, |k_z| \leq l_j/2\} \,,$$

with L_j, and l_j the smallest possible bounds, and t_j the largest possible bound, such that $\chi_{j,0}(\boldsymbol{k}) = 0$ outside this rectangle R. Here L_j and l_j determine, respectively, the sampling intervals (or translation steps) in x- and z-direction in the spatial domain. From a Fourier series expansion for functions with period $2L$, $e^{iq\pi x/L}/\sqrt{2L}$, with $q \in \mathbb{Z}$, is an orthonormal basis for L^2 functions with period $2L$. This means that $e^{i2\pi m_2 k_z/l_j}/\sqrt{l_j}$, with $m_2 \in \mathbb{Z}$, is an orthonormal basis for $L^2(I_{2j})$. Also, $e^{i\pi(m_1+1/2)k_x/L_j}/\sqrt{2L_j}$ with $m_1 \in \mathbb{Z}$, is an orthonormal basis for $L^2(\pm I_{1j})$ [see e.g. Hernández & Weiss (1996, Chapter 1.4)]. Therefore, the sequence $(u_{j,m})_{m\in\mathbb{Z}^2}$ [with $m = (m_1, m_2)$] defined as

$$u_{j,m}\left(k_x, k_z\right) = \frac{e^{i\pi(m_1+1/2)k_x/L_j} e^{i2\pi m_2 k_z/l_j}}{\sqrt{2L_j l_j}}$$

is an orthonormal basis for $L^2\left(\pm I_{1j} \times I_{2j}\right)$. Defining δ_1 and δ_2 by $L_j = \delta_1 \pi 2^j$ and $l_j = \delta_2 2\pi 2^{\lfloor j/2 \rfloor}$, we have

$$u_{j,m}\left(k_x, k_z\right) = \frac{2^{-3j/4}}{2\pi\sqrt{\delta_1\delta_2}} e^{i\pi(m_1+1/2)2^{-j}k_x/\delta_1} e^{im_2 2^{-\lfloor j/2 \rfloor}k_z/\delta_2} \,,$$

where the multiplicative term $2^{-3j/4}$ follows from ignoring the splitting at every other scale in the term $2^{-(\lfloor j/2 \rfloor + j)/2}$, i.e., replacing $\lfloor j/2 \rfloor$ with $j/2$ in this term.

Using this definition, and denoting by $R_{\theta_{j,l}}$ the rotation by angle $\theta_{j,l}$, the curvelet in

[1]This scaling relation holds when both width and length are made dimensionless through division by a reference length, typically the sampling interval in numerical implementations. Otherwise this relation would involve a proportionality constant that has a dimension which adjusts for the difference in dimension between width and squared length^2.

the frequency domain is then defined as

$$\hat{c}_\mu(\boldsymbol{k}) = 2\pi\chi_{j,l}(\boldsymbol{k})u_{j,m}\left(R^*_{\theta_{j,l}}\boldsymbol{k}\right), \quad \mu = (j,l,m), \tag{H.10}$$

where $\hat{c}_\mu$ denotes the Fourier transform of the curvelet $c_\mu \in L^2(\mathbb{R}^2)$. These are the *fine scale* curvelets. The coarse scale curvelets are then defined as $\hat{c}_{\mu_0}(\boldsymbol{k}) = 2\pi\chi_0(\boldsymbol{k})u_m(\boldsymbol{k})$, with $u_m(\boldsymbol{k}) = (2\pi\delta_0)^{-1}e^{i(m_1k_x/\delta_0 + m_2k_z/\delta_0)}$. Here, δ_0 is again some appropriate constant determining the sampling rate (or translations steps) in the spatial domain. Note that here δ_0 determines the sampling rate in both the x- and z-direction, since the coarse scale curvelets are isotropic (and thus really are wavelets).

Using the definition of curvelets given in equation (H.10), curvelets form a normalized tight frame. From equation (H.10), in L^2 we have

$$\sum_{m_1,m_2}\left|\left(\hat{f},\hat{c_\mu}\right)\right|^2 = (2\pi)^2\cdot\int\left|\hat{f}(\boldsymbol{k})\right|^2|\chi_{j,l}(\boldsymbol{k})|^2\,d\boldsymbol{k},$$

where we have used that $\left(u_{j,m}\left(R^*_{\theta_{j,l}}\boldsymbol{k}\right)\right)_{m\in\mathbb{Z}^2}$ is by construction an orthonormal basis for L^2 over the support of $\chi_{j,l}(\boldsymbol{k})$, i.e., $\left|u_{j,m}\left(R^*_{\theta_{j,l}}\boldsymbol{k}\right)\right|^2 = 1$. Using that the $\chi_{j,l}(\boldsymbol{k})$ constitute a partition of unity by equation (H.8), it then follows that

$$\begin{aligned}
\sum_{j\geq 0}\sum_{l=0}^{2^{\lfloor j/2\rfloor}-1}\sum_{m_1,m_2}\left|\left(\hat{f},\hat{c}_{(j,l,m_1,m_2)}\right)\right|^2 &= (2\pi)^2\sum_{j\geq 0}\sum_{l=0}^{2^{\lfloor j/2\rfloor}-1}\int\left|\hat{f}(\boldsymbol{k})\right|^2|\chi_{j,l}(\boldsymbol{k})|^2\,d\boldsymbol{k}\\
&= (2\pi)^2\int\left|\hat{f}(\boldsymbol{k})\right|^2\left[\sum_{j\geq 0}\sum_{l=0}^{2^{\lfloor j/2\rfloor}-1}|\chi_{j,l}(\boldsymbol{k})|^2\right]d\boldsymbol{k}\\
&= (2\pi)^2\int\left|\hat{f}(\boldsymbol{k})\right|^2 d\boldsymbol{k}\\
&= (2\pi)^2\left\|\hat{f}\right\|_2^2.
\end{aligned}$$

Therefore $(\hat{c}_\mu)_{\mu\in M}$ (with $\mu = (j,l,m_1,m_2)$ a multi-index) is a tight frame for $L^2(\mathbb{R}^2)$. Here, M is a multi-index set that has the appropriate ranges for j, l, m_1 and m_2. Finally, using the Parseval formula and the Plancherel formula for $L^2(\mathbb{R}^2)$, i.e., $\left(\hat{f},\hat{c}_\mu\right) = (2\pi)^2(f,c_\mu)$ and $\left\|\hat{f}\right\|_2^2 = (2\pi)^2\|f\|_2^2$, respectively, it follows that for $f\in L^2(\mathbb{R}^2)$

$$\sum_{\mu\in M}|(f,c_\mu)|^2 = \|f\|_2^2\,. \tag{H.11}$$

This shows that $(c_\mu)_{\mu\in M}$ is a normalized tight frame for $L^2(\mathbb{R}^2)$, giving the reconstruction formula (H.1).

Appendix I

Tight frames in Hilbert spaces

Let $\mathcal{H}$ be a Hilbert space, and let $(\cdot,\cdot)$ denote the inner product on $\mathcal{H}$. A collection of elements $\{e_j\}_{j\in J}$ in $\mathcal{H}$, with J an index set, is called a **frame** if there exist constants A and B, $0 < A \leq B < \infty$, such that

$$A\,\|f\|_{\mathcal{H}}^2 \leq \sum_{j\in J} |(f,e_j)|^2 \leq B\,\|f\|_{\mathcal{H}}^2 \quad \forall f \in \mathcal{H}, \tag{I.1}$$

where A and B are called **frame bounds**, and where $\|\cdot\|_{\mathcal{H}}$ denotes the norm on $\mathcal{H}$ induced by the inner product. If the frame bounds are equal ($A = B$), the frame is called **tight**. For example, the four vectors $\phi_1 = (0,1)$, $\phi_2 = (1,0)$, $\phi_3 = \left(\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2}\right)$, and $\phi_4 = \left(-\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2}\right)$, form a tight frame for $\mathbb{R}^2$ with frame bounds $A = B = 2$, since it follows that $\sum_{i=1}^4 |(f,\phi_i)|^2 = 2\,\|f\|^2$.

The linear span of frame elements is *dense* in $\mathcal{H}$ [e.g., Hernández & Weiss (1996, p.399)], meaning that any element $f \in \mathcal{H}$ can be written as a linear combination of the frame elements, and that the difference between this linear combination and f has a measure zero. If the frame elements were also linearly independent, they would form a basis for $\mathcal{H}$. The frame elements e_j are not a basis since adding the zero vector to $\{e_j\}_{j\in J}$ does not alter the inequalities in expression (I.1). Adding any vector, even a zero vector, to a basis, destroys the linear independence of the basis, meaning it no longer would be a basis. When both frame bounds equal unity ($A = B = 1$), the frame is called a *normalized* tight frame. Each orthonormal basis is a normalized tight frame, but the converse is not generally true, since the frame elements need not be linearly independent.

Although the frame elements do not need to be normal, they must satisfy $\|e_j\| \leq \sqrt{B}$, $\forall j \in J$. To see this, let $f = e_k$ for some $k \in J$, and use frame condition (I.1) to see that $\|e_k\|^4 = |(e_k,e_k)|^2 \leq \sum_{j\in J} |(e_k,e_j)|^2 \leq B\,\|e_k\|^2$. Therefore we have $\|e_k\| \leq \sqrt{B}$, $\forall k \in J$. If the frame is a normalized tight frame (i.e., $A = B = 1$), we must have $\|e_k\| \leq 1$. Hence, for a normalized tight frame, the linear independence of a basis is traded for the condition $\|e_k\| \leq 1$, $\forall k \in J$.

If a normalized tight frame has $\|e_k\| = 1$, $\forall k \in J$, $\{e_j\}_{j\in J}$ is an orthonormal basis for $\mathcal{H}$. To see this, observe that for fixed $k \in J$, $1 = \|e_k\|^4 = |(e_k,e_k)|^2 \leq \sum_{j\in J} |(e_k,e_j)|^2 \leq 1$. Therefore, we have $\sum_{j\in J} |(e_k,e_j)|^2 = |(e_k,e_k)|^2$

$+\sum_{j\neq k, j\in J}|(e_k,e_j)|^2 = \|e_k\|^2 + \sum_{j\neq k,j\in J}|(e_k,e_j)|^2 = 1$, and thus $\sum_{j\neq k,j\in J}|(e_k,e_j)|^2 = 0$. This implies $(e_k,e_j)=0$, $\forall k\in J$ and $k\neq j$. Since k is arbitrary, and since $\|e_k\|=1$, it follows that $\{e_j\}_{j\in J}$ is an orthonormal basis for $\mathcal{H}$.

It is known [e.g., Hernández & Weiss (1996, pp.334-336)] that for a Hilbert space $\mathcal{H}$ and a family of elements $\{e_j\}_{j\in J}$ in $\mathcal{H}$, the condition

$$\|f\|_{\mathcal{H}}^2 = \sum_{j\in J}|(f,e_j)|^2 \quad \forall f\in\mathcal{H}, \tag{I.2}$$

is necessary and sufficient for

$$f = \sum_{j\in J}(f,e_j)\,e_j \tag{I.3}$$

to hold. That is, statements (I.2) and (I.3) are equivalent. This tells us that for a tight frame with frame bound A, reconstruction formula (I.3) holds, for if $\sum_{j\in J}|(f,e_j)|^2 = A\|f\|_{\mathcal{H}}^2$, simply defining $e'_j = e_j/\sqrt{A}$, gives $\sum_{j\in J}\left|\left(f,e'_j\right)\right|^2 = \|f\|_{\mathcal{H}}^2$. From the equivalence of statements (I.2) and (I.3), then $f = \sum_{j\in J}\left(f,e'_j\right)e'_j = \frac{1}{A}\sum_{j\in J}(f,e_j)\,e_j$; i.e., for a tight frame we have reconstruction formula (I.3). For a normalized tight frame ($A=1$), $f = \sum_{j\in J}(f,e_j)\,e_j$, and thus also the reconstruction formula (I.3). Note that this reconstruction formula is identical to the reconstruction formula for an orthonormal basis, but here the frame elements are not orthogonal; i.e., you can have reconstruction formula (I.3) with linearly dependent elements of $\mathcal{H}$, provided (I.2) holds.

Although we showed that the reconstruction formula (I.3) holds for (normalized) tight frames, a similar reconstruction formula can be found if the frame is not tight. We refer the reader to Mallat (1998, chapter 5), Hernández & Weiss (1996, Chapter 8), or Strang & Nguyen (1997, Section 2.5) for treatments of non-tight frames.

Appendix J

The quartic equation for the stationary midpoint location as a function of y and p

The stationary midpoint location $x(\boldsymbol{y}, p; h)$ satisfies the condition that the derivative of the phase Φ with respect to midpoint location is zero. That is, using equation (4.11), $x(\boldsymbol{y}, p; h)$ satisfies

$$\left(\frac{\partial}{\partial x}\left[\frac{(r_s + r_g)}{v} - px\right]\right) = 0\,, \tag{J.1}$$

with r_s and r_g given in equation (4.9). Solving this equation for $x(\boldsymbol{y}, p; h)$ leads to a quartic equation in $\sigma := (x(\boldsymbol{y}, p; h) - y_1)^2$ given by

$$\sigma^4 + a_3\sigma^3 + a_2\sigma^2 + a_1\sigma + a_0 = 0\,, \tag{J.2}$$

with

$$a_3 := \frac{y_2^2\left(3 - (vp)^2\right)}{(1 - (vp/2)^2)} - 4h^2\,, \tag{J.3}$$

$$a_2 := 6h^4 - \frac{h^2y_2^2\left(3 - (vp)^2\right) - \frac{6}{4}y_2^4\left(2 - (vp)^2\right)}{1 - (vp/2)^2}\,, \tag{J.4}$$

$$a_1 := \frac{y_2^6\left(1 - (vp)^2\right) - h^4y_2^2\left(3 - (vp)^2\right) - h^2v^2y_2^4\left(4/p^2 - 2v^2 + p^2v^4\right)}{1 - (vp/2)^2} - 4h^6, \tag{J.5}$$

$$a_0 := \left(h^2 + y_2^2\right)^3\left[h^2 - \frac{(vp)^2y_2^2}{4 - (vp)^2}\right]\,. \tag{J.6}$$

Note that for $p > 0$ we have $y_1 < x$, and that for $p < 0$ we have $y_1 > x$. Therefore, in general we have

$$x(\boldsymbol{y}, p; h) = \operatorname{sgn}(p)\sqrt{\sigma} + y_1\,. \tag{J.7}$$

The general solution to a quartic equation is given in Appendix G.

Appendix K

Explicit expression for the linear transformation T for common-offset time-migration with curvelets

To find an explicit expression for the linear transformation T given in equation (4.37), first consider the term $\nabla_{\boldsymbol{y}} t(\boldsymbol{y}, x(\boldsymbol{y}, p_u; h); h)|_{\boldsymbol{y}_m}$. For the subsequent derivation it is important to realize that in calculating the gradient of t with respect to $\boldsymbol{y}$, the midpoint location x is fixed. Using equation (4.19) together with the definitions of r_s and r_g in equation (4.9), it follows that

$$\begin{aligned}\nabla_{\boldsymbol{y}} t(\boldsymbol{y}, x(\boldsymbol{y}, p_u; h); h)|_{\boldsymbol{y}_m} &= \frac{1}{v}\begin{pmatrix} -\left(\dfrac{x_u - h - y_1^m}{r_s} + \dfrac{x_u + h - y_1^m}{r_g}\right) \\ \\ \dfrac{y_2^m}{r_s} + \dfrac{y_2^m}{r_g} \end{pmatrix} \\ &= \frac{1}{v}\begin{pmatrix} -(\sin\theta_s + \sin\theta_g) \\ \\ \cos\theta_s + \cos\theta_g \end{pmatrix} , \end{aligned} \tag{K.1}$$

where $\boldsymbol{y}_m = (y_1^m \; y_2^m)^T$, and where the angles θ_s and θ_g are defined in Figure 4.7. Then, defining the slowness vectors $\boldsymbol{p}_s$ and $\boldsymbol{p}_g$ at the source and receiver locations, respectively, as

$$\boldsymbol{p}_{s,g} := \frac{1}{v}\begin{pmatrix} -\sin\theta_{s,g} \\ \cos\theta_{s,g} \end{pmatrix} , \tag{K.2}$$

we can rewrite equation (K.1) as

$$\nabla_{\boldsymbol{y}} t(\boldsymbol{y}, x(\boldsymbol{y}, p_u; h); h)|_{\boldsymbol{y}_m} = \boldsymbol{p}_s + \boldsymbol{p}_g := \boldsymbol{p}_m , \tag{K.3}$$

Here, $\boldsymbol{p}_m$ is the slowness vector associated with the dip covector $\boldsymbol{\xi}_m := \omega \boldsymbol{p}_m$ (i.e., the wave vector associated with the reflector). Defining $\phi := (\theta_s + \theta_g)/2$ as the dip (i.e., the angle with the horizontal measured clockwise positive), it follows that

$$\nabla_{\boldsymbol{y}} t(\boldsymbol{y}, x(\boldsymbol{y}, p_u; h); h)|_{\boldsymbol{y}_m} = \frac{2\cos\theta}{v}\begin{pmatrix} -\sin\phi \\ \cos\phi \end{pmatrix} , \tag{K.4}$$

where $\theta := (\theta_g - \theta_s)/2$ is the half opening-angle.

Next, consider the term $\nabla_{\boldsymbol{y}} x(\boldsymbol{y}, p_u; h)|_{\boldsymbol{y}_m}$. From Appendix J, the stationary midpoint

location satisfies equation (J.1). Explicitly writing out this equation gives

$$\frac{x-h-y_1}{r_s}+\frac{x+h-y_1}{r_g}+pv=0\,. \tag{K.5}$$

Now, treating x as the dependent variable and $\boldsymbol{y}$ and $p=k_x/\omega$ as the independent variables, we can take the partial derivative with respect to y_1 on both sides of equation (K.5) to get

$$\left(\frac{\partial x}{\partial y_1}-1\right)\left\{\left(\frac{1}{r_s}+\frac{1}{r_g}\right)+\frac{(x-h-y_1)^2}{r_s^3}+\frac{(x+h-y_1)^2}{r_g^3}\right\}=0\,. \tag{K.6}$$

Therefore, we have

$$\left.\frac{\partial x(\boldsymbol{y},p;h)}{\partial y_1}\right|_{\boldsymbol{y}_m}=1\,. \tag{K.7}$$

This result is intuitive, since changing the horizontal component y_1 of the image location $\boldsymbol{y}$, while keeping the slope p constant, should result in a simple translation of the whole geometry shown in Figure 4.7 along the horizontal axis. Similarly, taking the partial derivative with respect to y_2 on both sides of equation (K.5), we get

$$\frac{1}{r_s}\frac{\partial x}{\partial y_2}-\frac{x-h-y_1}{r_s^2}\frac{\partial r_s}{\partial y_2}+\frac{1}{r_g}\frac{\partial x}{\partial y_2}-\frac{x+h-y_1}{r_g^2}\frac{\partial r_g}{\partial y_2}=0\,. \tag{K.8}$$

Using that

$$\frac{\partial r_s}{\partial y_2}=\frac{x-h-y_1}{r_s^2}\frac{\partial x}{\partial y_2}+\frac{y_2}{r_s}\,, \tag{K.9}$$

$$\frac{\partial r_g}{\partial y_2}=\frac{x+h-y_1}{r_g^2}\frac{\partial x}{\partial y_2}+\frac{y_2}{r_g}\,, \tag{K.10}$$

and substituting these expressions in equation (K.8), it follows that

$$\left.\frac{\partial x(\boldsymbol{y},p)}{\partial y_2}\right|_{\boldsymbol{y}_m}=\frac{\tan\theta_s r_g^3+\tan\theta_g r_s^3}{r_g^3+r_s^3}\,. \tag{K.11}$$

Using this together with equation (K.7), it follows that

$$\nabla_{\boldsymbol{y}}x(\boldsymbol{y},p_u;h)\big|_{\boldsymbol{y}_m}=\begin{pmatrix}1 & \dfrac{\tan\theta_s r_g^3+\tan\theta_g r_s^3}{r_g^3+r_s^3}\end{pmatrix}^T\,. \tag{K.12}$$

Finally, using the results from equations (K.4) and (K.12) in equation (4.37), we find the

resulting explicit expression for the linear transformation T to be

$$T = \begin{pmatrix} 1 & \dfrac{\tan\theta_s r_g^3 + \tan\theta_g r_s^3}{r_s^3 + r_g^3} \\ \dfrac{-2\cos\theta\sin\phi}{v} & \dfrac{2\cos\theta\cos\phi}{v} \end{pmatrix}. \tag{K.13}$$

Appendix L

Discrete form of the linear transformation L^{-1}

For the purpose of digitally implementing the linear transformation L^{-1}, we express the shear, rotation, and dilation matrices that together form L as matrices on a grid with sampling intervals Δt and Δx in the time and midpoint axes, respectively. Starting with the shear matrix S_{-p_u} [see equation (4.36)], we observe that

$$S_{-p_u} \cdot \begin{pmatrix} x \\ t \end{pmatrix} = \begin{pmatrix} x \\ t - p_u x \end{pmatrix} . \tag{L.1}$$

Writing the vector $(x \quad t)^T$ as $(n_x \Delta x \quad n_t \Delta t)^T$, it follows that

$$S_{-p_u} \cdot \begin{pmatrix} x \\ t \end{pmatrix} = \begin{pmatrix} n_x \Delta x \\ (n_t - p_u \dfrac{\Delta x}{\Delta t} n_x) \Delta t \end{pmatrix} . \tag{L.2}$$

Therefore, the matrix S_{-pu} expressed on the discrete grid, with sampling intervals Δt and Δx, is given by

$$S^{\Delta}_{-p_u} = \begin{pmatrix} 1 & 0 \\ -p_u \dfrac{\Delta x}{\Delta t} & 1 \end{pmatrix} , \tag{L.3}$$

where we use the superscript Δ to denote that the matrix is defined on a discrete grid. Hence, defining the vector $\boldsymbol{x}^{\Delta} := (n_x \quad n_t)^T$, we have

$$S^{\Delta}_{-p_u} \cdot \boldsymbol{x}^{\Delta} = \begin{pmatrix} n_x \\ n_t - p_u \dfrac{\Delta x}{\Delta t} n_x \end{pmatrix} . \tag{L.4}$$

Going through the same analysis for the matrices D^{-1} and S^{-1} [equation (4.46)], we find

$$\left(D^{\Delta}\right)^{-1} = \begin{pmatrix} \dfrac{1}{\cos\phi + X \sin\phi} & 0 \\ 0 & \dfrac{1}{\cos\theta} \end{pmatrix} , \tag{L.5}$$

$$\left(S^{\Delta}\right)^{-1} = \begin{pmatrix} 1 & -\dfrac{X\cos\phi - \sin\phi}{\cos\phi + X\sin\phi} \dfrac{v\Delta t}{2\Delta x} \\ 0 & 1 \end{pmatrix} , \tag{L.6}$$

while for the rotation matrix $R_{-\phi}$ we get

$$R_{-\phi}^{\Delta} = \begin{pmatrix} \cos\phi & -\sin\phi\dfrac{v\Delta t}{2\Delta x} \\ \sin\phi\dfrac{2\Delta x}{v\Delta t} & \cos\phi \end{pmatrix} . \tag{L.7}$$

In finding the expressions for $\left(D^{\Delta}\right)^{-1}$, $\left(S^{\Delta}\right)^{-1}$, and $R_{-\phi}^{\Delta}$, we used that the sampling interval Δz in the image satisfies $\Delta z = v\Delta t/2$.

Therefore, using equations (L.4)–(L.7), we have

$$\begin{pmatrix} n_x' \\ n_z \end{pmatrix} = \left(L^{\Delta}\right)^{-1} \cdot \begin{pmatrix} n_x \\ n_t \end{pmatrix} = R_{-\phi}^{\Delta} \cdot \left(S^{\Delta}\right)^{-1} \cdot \left(D^{\Delta}\right)^{-1} \cdot S_{-p_u}^{\Delta} \cdot \begin{pmatrix} n_x \\ n_t \end{pmatrix} . \tag{L.8}$$

In general, the new grid coordinates n_x' and n_z are not integers. Therefore we need to interpolate the obtained grid coordinates to integer values of n_x' and n_z. We do this using bilinear interpolation.

Appendix M

[2/2] rational interpolation for nonhyperbolic moveout in a single horizontal VTI layer

Using the definition and normalization of the rational approximation outlined in the main text, we can write the [2/2] rational approximation for squared traveltimes T as a function of squared offset X as

$$T(X) \approx \frac{T_0 + n_1 X + n_2 X^2}{1 + d_1 X + d_2 X^2} \quad , \tag{M.1}$$

where $T_0 = t_0^2$ is the squared zero-offset two-way traveltime, and $n_{1,2}$ and $d_{1,2}$ are the coefficients of the numerator and denominator of the rational approximant, respectively. Using four squared traveltimes $T_i = t_i^2$, with $i = 1, \ldots, 4$, and four accompanying squared offsets $X_i = x_i^2$ as interpolation points, we arrive at a linear system of four equations with four unknowns, the coefficients $n_{1,2}$ and $d_{1,2}$. We simply solve (using Mathematica) for the four coefficients in terms of T_i, X_i, and T_0. The resulting expressions for the coefficients are given by

$$\begin{aligned}
d_1 &= ((T_0 - T_4)X_1X_2X_3(T_1X_1(X_2 - X_3) + T_3X_3(X_1 - X_2) + \ldots \\
&\quad T_2X_2(X_3 - X_1)) - ((T_1 - T_2)(T_0 - T_3)X_1^2X_2^2 + ((T_2 - T_0)(T_1 - T_3)X_1^2 + \ldots \\
&\quad (T_0 - T_1)(T_2 - T_3)X_2^2)X_3^2)X_4 + ((T_0 - T_3)X_1X_2(T_1X_1 - T_2X_2 + \ldots \\
&\quad T_4(X_2 - X_1)) - ((T_0 - T_2)(T_1 - T_4)X_1^2 - (T_0 - T_1)(T_2 - T_4)X_2^2)X_3 + \ldots \\
&\quad (T_3 - T_4)(T_1X_2 - T_2X_1 + T_0(X_1 - X_2))X_3^2)X_4^2)/\ldots \\
&\quad (X_1X_2X_3(T_2(T_3(X_2 - X_3)(X_1 - X_4) - T_4(X_1 - X_3)(X_2 - X_4)) + \ldots \\
&\quad T_1(T_4(X_2 - X_3)(X_1 - X_4) - T_3(X_1 - X_3)(X_2 - X_4) + \ldots \\
&\quad T_2(X_1 - X_2)(X_3 - X_4)) + T_3T_4(X_1 - X_2)(X_3 - X_4))X_4) \, , \qquad \text{(M.2)} \\
d_2 &= (-T_0(X_1 - X_2)(X_1 - X_3)(X_2 - X_3) + T_3X_1X_2(X_1 - X_2)(1 + d_1X_3) + \ldots \\
&\quad X_3(T_1X_2(1 + d_1X_1)(X_2 - X_3) + T_2X_1(1 + d_1X_2)(X_3 - X_1)))/\ldots \\
&\quad (X_1X_2X_3(T_2X_2(X_1 - X_3) + T_3X_3(X_2 - X_1) + T_1X_1(X_3 - X_2))) \, , \qquad \text{(M.3)} \\
n_2 &= T_0(X_1 - X_2) + T_1(1 + X_1(d_1 + d_2X_1))X_2 - T_2X_1(1 + X_2(d_1 + d_2X_2))/\ldots \\
&\quad X_1X_2(X_1 - X_2) \, , \qquad \text{(M.4)} \\
n_1 &= T_1d_1 - \frac{T_0 - T_1}{X_1} - (n_2 - d_2T_1)X_1 \, . \qquad \text{(M.5)}
\end{aligned}$$

Note that only the expression for d_1 is explicit in just T_i, X_i, and T_0, whereas d_2 also depends on d_1, n_2 on $d_{1,2}$, and n_1 on n_2 and $d_{1,2}$. Calculating the coefficients $n_{1,2}$ and $d_{1,2}$ using the above expressions and the interpolation traveltimes t_i and offsets x_i, we can use the resulting values of $n_{1,2}$ and $d_{1,2}$ in equation (M.1) to evaluate interpolated traveltimes t for offsets x between offsets x_i.

Appendix N

Degeneracy of [2/2] rational interpolation for hyperbolic moveout

Replacing the traveltime t and offset x in equation (5.14) by the squared traveltimes T and squared offsets X and adopting the normalization $N_L(0) = T(0)$ and $D_M(0) = 1.0$ after Baker (1975, pp. 5-6), the [2/2] rational interpolation is given by

$$T(X) \approx \frac{T_0 + n_1 X + n_2 X^2}{1 + d_1 X + d_2 X^2} \quad . \tag{N.1}$$

Here T_0 denotes the squared zero-offset two-way traveltime. Using four support points (X_i, T_i), we arrive at a linear system of equations for the unknown interpolation coefficients n_1, n_2, d_1, d_2, i.e.,

$$A \cdot \boldsymbol{x} = \boldsymbol{d} \, . \tag{N.2}$$

where

$$A := \begin{pmatrix} -X_1 & -X_1^2 & X_1 T_1 & X_1^2 T_1 \\ -X_2 & -X_2^2 & X_2 T_2 & X_2^2 T_2 \\ -X_3 & -X_3^2 & X_3 T_3 & X_3^2 T_3 \\ -X_4 & -X_4^2 & X_4 T_4 & X_4^2 T_4 \end{pmatrix}, \; \boldsymbol{x} := \begin{pmatrix} n_1 \\ n_2 \\ d_1 \\ d_2 \end{pmatrix}, \; \boldsymbol{d} := \begin{pmatrix} T_0 - T_1 \\ T_0 - T_2 \\ T_0 - T_3 \\ T_0 - T_4 \end{pmatrix}, \tag{N.3}$$

with T_0 the squared two-way zero-offset traveltime. Calculating $\det A$, we get

$$\begin{aligned} \det A \;\; = \;\; & X_1 X_2 X_3 X_4 \left[(T_1 T_2 + T_3 T_4)(X_1 - X_2)(X_3 - X_4) \ldots \right. \\ & + (T_1 T_2 + T_3 T_4)(X_1 - X_2)(X_3 - X_4) \ldots \\ & \left. - (T_1 T_2 + T_3 T_4)(X_1 - X_2)(X_3 - X_4) \right] \, . \end{aligned} \tag{N.4}$$

When the moveout is hyperbolic, we have $T_i = T_0 + X_i / v^2$. Using this, it follows that

$$X_i - X_j = v^2 (T_i - T_j) \, . \tag{N.5}$$

Using equation (N.5) in (N.4) then gives

$$\begin{aligned}
\det A &= v^2 X_1 X_2 X_3 X_4 \left[(T_1T_2 + T_3T_4)(T_1T_3 - T_1T_4 - T_2T_3 + T_2T_4) \dots \right. \\
&\quad + (T_1T_2 + T_3T_4)(T_1T_2 - T_2T_4 - T_1T_3 + T_3T_4) \dots \\
&\quad \left. - (T_1T_2 + T_3T_4)(T_1T_2 - T_1T_4 - T_2T_3 + T_3T_4) \right] \\
&= v^2 X_1 X_2 X_3 X_4 \left[(T_1T_2 + T_3T_4)(T_1T_3 - T_1T_4 - T_2T_3 + T_2T_4) \dots \right. \\
&\quad - (T_1T_2 + T_3T_4)(T_1T_3 - T_1T_4 - T_2T_3 + T_2T_4) \dots \\
&\quad \left. + (T_1T_3 + T_2T_4)(T_1T_4 + T_2T_3) - (T_1T_3 + T_2T_4)(T_1T_4 + T_2T_3) \right] \\
&= 0 . \qquad \text{(N.6)}
\end{aligned}$$

Hence, for hyperbolic moveout, the linear system (N.2) for [2/2] rational interpolation of the moveout (expressed as squared traveltime as a function of squared offset) becomes degenerate.

Appendix O

A condition to estimate the reliability of the zero-lag cross-correlation correction factor

Equation (7.16) for the correction factor of the cross-correlation coefficient depends on the unknown noise functions $n_u(t)$ and $n_p(t)$. In practice, we don't know these noise functions and often have only an estimate of the noise level, as opposed to some ensemble average. We want the correction factor to be reliable when the unknown variances of the noise in equation (7.16) are replaced by a single estimate of the noise level. If inequalities (7.17)-(7.19) are satisfied, the correction factor depends only weakly on the estimate of the noise level. Here we rewrite these inequalities into a single inequality that can be verified using the noise-contaminated signals and an estimate of a single realization of the noise $n_0(t)$.

To write inequalities (7.17)-(7.19) as a single one, we first add the left- and right-hand sides of equations (7.17)-(7.19), while multiplying equation (7.17) by two for convenience in the further derivation. This gives

$$\begin{aligned} 2\left(|(u_u+n_u,n_p)_{(t_s;t,t_w)}+(u_p,n_u)_{(t_s;t,t_w)}|+|(u_u,n_u)_{(t_s=0;t,t_w)}|+|(u_p,n_p)_{(t_s=0;t,t_w)}|\right) \\ \ll \\ 2|(u_u,u_p)_{(t_s;t,t_w)}|+(u_u,u_u)_{(t_s=0;t,t_w)}+(n_u,n_u)_{(t_s=0;t,t_w)} \\ +(u_p,u_p)_{(t_s=0;t,t_w)}+(n_p,n_p)_{(t_s=0;t,t_w)}\,, \end{aligned} \tag{O.1}$$

where we have used that zero-lag auto-correlations are positive definite. If conditions (7.17)-(7.19) hold, equations (7.13)-(7.15) from the main text are good approximations. We can use these approximations, together with the linearity of $(\cdot\,,\cdot)_{(t_s;t,t_w)}$ and equation (7.5), to approximate inequality (O.1) as

$$\begin{aligned} 2|(u'_u,n_p)_{(t_s;t,t_w)}+(u'_p,n_u)_{(t_s;t,t_w)}-(n_u,n_p)_{(t_s;t,t_w)}| \\ +2|(u'_u,n_u)_{(t_s=0;t,t_w)}-(n_u,n_u)_{(t_s=0;t,t_w)}|+2|(u'_p,n_p)_{(t_s=0;t,t_w)}-(n_p,n_p)_{(t_s=0;t,t_w)}| \\ \ll \\ 2|(u'_u,u'_p)_{(t_s;t,t_w)}|+(u'_u,u'_u)_{(t_s=0;t,t_w)}+(u'_p,u'_p)_{(t_s=0;t,t_w)}\,. \end{aligned} \tag{O.2}$$

For the applications treated in this paper, i.e., a source displacement and the displacement of the scatterer locations, the maximum cross-correlation occurs for $t_s = \langle\tau\rangle = 0$ s. Also, since in CWI we assume that the traveltime perturbations in the time window $[t-t_w, t+t_w]$ are small, we expect a positive cross-correlation between the signals $u'_u(t)$ and

$u'_p(t)$ for times where CWI is valid. Hence, making use of this approximation, and setting $t_s = 0$ in condition (O.2), we can write this condition as

$$\begin{aligned}|(u'_u + u'_p, n_0)_{(t_s=0;t,t_w)}| + |(u'_u, n_0)_{(t_s=0;t,t_w)} - (n_0, n_0)_{(t_s=0;t,t_w)}| \\ +|(u'_p, n_0)_{(t_s=0;t,t_w)} - (n_0, n_0)_{(t_s=0;t,t_w)}| \ll 2([u'_u + u'_p]/2, [u'_u + u'_p]/2)_{(t_s=0;t,t_w)}\end{aligned} \tag{O.3}$$

where we substituted for both noise signals $n_u(t)$ and $n_p(t)$ the estimated noise signal $n_0(t)$, and we assumed $\left|(n_u, n_p)_{(t_s;t,t_w)}\right| \ll \left|(u'_u, n_p)_{(t_s;t,t_w)} + (u'_p, n_u)_{(t_s;t,t_w)}\right|$ to eliminate the $(n_u, n_p)_{(t_s;t,t_w)}$ term. The latter approximation is more appropriate for larger signal-to-noise ratios, larger time windows, and uncorrelated noise realizations $n_u(t)$ and $n_p(t)$. Note that substituting a single noise signal for both unknown noise signals $n_u(t)$ and $n_p(t)$ is appropriate only if both noise signals have similar noise levels. Dividing both sides of condition (O.3) by $(n_0, n_0)_{(t_s=0;t,t_w)}$ (which is positive definite), and defining the ratio

$$\Gamma(t, t_w) := \frac{\left([u'_u + u'_p]/2, [u'_u + u'_p]/2\right)_{(t_s=0;t,t_w)}}{(n_0, n_0)_{(t_s=0;t,t_w)}}, \tag{O.4}$$

inequality (O.3) leads to

$$\begin{aligned}\frac{1}{2}\left(\left|\frac{(u'_u + u'_p, n_0)_{(t_s=0;t,t_w)}}{(n_0, n_0)_{(t_s=0;t,t_w)}}\right| + \left|\frac{(u'_u, n_0)_{(t_s=0;t,t_w)}}{(n_0, n_0)_{(t_s=0;t,t_w)}} - 1\right| + \left|\frac{(u'_p, n_0)_{(t_s=0;t,t_w)}}{(n_0, n_0)_{(t_s=0;t,t_w)}} - 1\right|\right) \\ /\Gamma(t, t_w) \leq \gamma,\end{aligned} \tag{O.5}$$

where γ is $O(10^{-1})$. Here $\sqrt{\Gamma(t, t_w)}$ can be interpreted as the average signal-to-noise ratio.

Condition (O.5) is satisfied only for time windows that have a large average signal-to-noise ratio. The l.h.s of condition (O.5) can be evaluated using the noise-contaminated signals $u'_u(t)$ and $u_p(t)$, and an estimated noise signal $n_0(t)$. This condition can thus be used as a selection criterion to determine which time windows have reliable correction factors associated with them, when an estimated noise level (calculated using an estimated noise realization $n_0(t)$) is used, i.e., when equation (7.20) is used to calculate the correction factor. Note that γ can be interpreted as the inverse of the signal-to-noise ratio.