On the modeling and inversion of seismic data

Over het modelleren en inverteren van seismische meetgegevens

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de Rector Magnificus, Prof. dr. H.O. Voorma, ingevolge het besluit van het College voor Promoties in het openbaar te verdedigen op maandag 11 december 2000 des namiddags te 16.15 uur

door

Christiaan Cornelis Stolk

geboren op 24 december 1973, te Bunnik

Promotor: Prof. dr. J.J. Duistermaat

Copromotor: Dr. A.P.E. ten Kroode
Shell Research, Rijswijk

Het in dit proefschrift beschreven onderzoek is gefinancierd door Shell International Exploration and Production BV

Stolk, Christiaan Cornelis
On the modeling and inversion of seismic data
Proefschrift Universiteit Utrecht – Met samenvatting in het Nederlands.

ISBN 90-393-2551-0

Aan mijn ouders

Contents

1 Introduction **1**
1.1 High-frequency asymptotics and Fourier integral operators 3
1.2 The seismic inversion problem 6
1.3 Contents of this thesis 10
1.3.1 Wave equations with discontinuous coefficients 10
1.3.2 Modeling and inversion of seismic data in anisotropic elastic media 11
1.3.3 Linearized inversion when traveltime injectivity is violated . . 12

2 Wave equations with discontinuous coefficients **15**
2.1 Introduction 15
2.2 The partial differential equation 16
2.3 The PDE on domains with a boundary 19
2.4 Well-posedness for general second order hyperbolic equations 22
2.5 More general differential equations 30
2.6 Solutions in larger or smaller spaces 37
2.7 Discontinuities at an interface 39
2.8 Continuous dependence on the coefficients 46
2.9 Derivative with respect to coefficients 52

3 Modeling and inversion of seismic data in anisotropic elastic media **55**
3.1 Introduction 55
3.2 Propagation of elastic waves in smoothly varying media 60
3.2.1 Decoupling the modes 60
3.2.2 The Green's function 63
3.3 Reflection at an interface: Microlocal analysis of the 'Kirchhoff' approximation 68
3.4 The Born approximation 74
3.4.1 Modeling: Perturbation of the Green's function 75
3.4.2 Restriction 77
3.4.3 Inversion 80
3.5 Symplectic geometry of the data 85
3.6 Modeling: Joint formulation 87
3.7 Inverse scattering revisited 89

3.7.1 Invertible transformation into subsurface coordinates 89
3.7.2 The inversion operator 90
3.7.3 Annihilators of the data 91

4 Linearized inversion when traveltime injectivity is violated 95
4.1 Introduction 95
4.2 Preliminaries, the linearized forward operator 99
4.3 Conditions for the normal operator to be a FIO 101
4.4 Sobolev estimates for the nonmicrolocal part and invertibility 107
4.5 Examples 111
4.6 Generically the normal operator is invertible 113
4.7 Discussion 119
4.A Perturbations of a ray due to medium perturbations 120

Bibliography 125

Samenvatting 129

Dankwoord 131

Curriculum Vitae 133

Chapter 1

Introduction

In a seismic experiment one tries to obtain information about the subsurface from measurements at the surface, using elastic waves. To this end elastic waves in the subsurface are generated by sources at the surface. The waves that return to the surface are observed. We will be interested in reflection seismology, where one uses the waves that are reflected in the subsurface. From these data one tries to reconstruct an image of the subsurface. This image can be interpreted to say more about the geological structure of the subsurface, for instance to aid in petroleum exploration.

Seismic experiments are conducted both on land and at sea. On land the source of waves may be an explosion or a heavy vibrating object (a vibroseis truck). The receivers (geophones) are usually placed along a line (to reconstruct a two-dimensional section of the subsurface), or on a grid (to reconstruct a three-dimensional part of the subsurface). They may record one component or three components of the movement of the surface. At sea the source is usually attached to a vessel, while the receivers are along one or more cables towed by the vessel.

To model the propagation of waves the subsurface is viewed as an acoustic or, more realistically, as an elastic medium[1]. Such a medium is described by a number of parameters that depend on the position. In the acoustic case these can be chosen for instance as the local wavespeed $c(x)$ and the mass density. In the elastic case these are the elastic tensor, and the mass density. The medium parameters occur as coefficients in the partial differential equations that describe the propagation of waves in the medium.

The waves generated by the source travel downward into the medium. At positions where the coefficients vary strongly part of the signal is reflected, and part is transmitted. The reflected signal travels upward and is observed when it hits the receivers. Often the medium consists of several layers, such that the medium properties vary little within a layer, but strongly at the interfaces between the layers. In such cases one can recognize in the data, after some preprocessing, arrivals of the signal

[1]For the theory of elasticity see Marsden and Hughes [37], or a standard textbook such as Landau and Lifshitz [35]. Of course this is only an approximation and many effects are not taken into account, such as dissipation, the presence of cracks, porosity of the rocks and the presence of fluids (water, oil, natural gas) in the pores.

reflected at different depths.

We view the subsurface as an elastic or acoustic medium with possibly discontinuous coefficients. Therefore the first question that we ask ourselves is whether in such media the solutions to the wave equation are still well defined. It turns out that this is known, and that indeed unique solutions can be found that depend in a stable manner on the source and the initial conditions. They also depend continuously on the coefficients, when the latter converge in a suitable sense. This is, with some other results, described in Chapter 2. Note that this theory does not say very much about the precise form of the solutions. Propagation of waves in media with singularities is still not well understood.

So the question is now whether the coefficients of an acoustic or elastic wave equation can be reconstructed from information the seismic data, that is, from information about the solutions at the boundary. This is an example of an inverse problem. Whereas the forward or direct problem leads to a problem that is well-posed in the sense of Hadamard, i.e. it has a unique solution that depends in a stable manner on the parameters, it is not clear whether this is the case for the inverse problem. Introductory texts about inverse problems are for instance the books by Kirsch [32] and Isakov [29]. More information about conventional processing techniques can be found in the books by Aki and Richards [2, 3], Sheriff and Geldart [49], and Yilmaz [60].

In Chapters 3 and 4 of this thesis we will study the inverse problem under additional assumptions, such that high-frequency methods can be used. The high frequencies in the Fourier decomposition correspond to the singularities of a function. High-frequency methods give very precise information about the solutions of wave equations in smoothly varying media. To apply such methods to media with singularities, we either assume that the singularities are jump discontinuities along smooth interfaces, or we use a linearization, the Born approximation introduced below.

In Chapter 3 we study what information about elastic media can be obtained from the high-frequency part of the data. We give detailed results about the reconstruction of the 'high-frequency part' of the medium coefficients, and indicate how it can be used to reconstruct the 'low frequency part' (the smooth part) of the medium. In Chapter 4 we study a more subtle problem for the Born approximation in acoustic media. The analysis is performed using microlocal analysis, that gives a very precise description of the singular (high-frequency) part of a function. In particular we use the theory of Fourier integral operators. In seismic data the low frequencies are absent, and in practice high-frequency methods are commonly applied to all data.

In Section 1.1 we give a short introduction to asymptotic solutions of the wave equation, microlocal analysis and Fourier integral operators. In Section 1.2 we describe some important ideas in the modeling and inversion of seismic data for the case of a relatively simple acoustic medium. In the last section of this introduction we describe in more detail the contents of Chapters 2 to 4.

1.1 High-frequency asymptotics and Fourier integral operators

The high-frequency part of solutions to wave equations can be found using microlocal analysis and the theory of Fourier integral operators. Other operators that occur in the analysis of seismic data are also of this type. For this reason we will describe some of the ideas behind this theory.

As an example we construct asymptotic solutions to the acoustic equation. These solutions will also be used in the next section. The acoustic equation for a medium of constant density is given by

$$Pu := \left(\frac{1}{c(x)^2}\frac{\partial^2}{\partial t^2} - \Delta\right) u = f. \tag{1.1}$$

In this section we assume that $c(x)$ is smooth. Let us try a solution of the form

$$v(x,t) = A(x,t,\tau)\mathrm{e}^{\mathrm{i}\tau\psi(x,t)}, \tag{1.2}$$

where $A(x,t,\tau) = \sum_{k\geq 0} A_k(x,t)\tau^{m-k}$ is the amplitude and $\mathrm{e}^{\mathrm{i}\tau\psi(x,t)}$ is an oscillatory factor with frequency parameter τ. We are interested in the case $\tau \to \infty$. We find

$$\begin{aligned}
Pv = {} & (\mathrm{i}\tau)^2 \left(\frac{1}{c^2}(\frac{\partial\psi}{\partial t})^2 - \sum_i (\frac{\partial\psi}{\partial x_i})^2\right) A\mathrm{e}^{\mathrm{i}\tau\psi} \\
& + \mathrm{i}\tau \left(2\frac{1}{c^2}\frac{\partial\psi}{\partial t}\frac{\partial A}{\partial t} - 2\sum_i \frac{\partial\psi}{\partial x_i}\frac{\partial A}{\partial x_i} + \frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2}A - \sum_i \frac{\partial^2\psi}{\partial x_i^2}A\right)\mathrm{e}^{\mathrm{i}\tau\psi} \\
& + \left(\frac{1}{c^2}\frac{\partial^2 A}{\partial t^2} - \sum_i \frac{\partial^2 A}{\partial x_i^2}\right)\mathrm{e}^{\mathrm{i}\tau\psi}.
\end{aligned}$$

We want this to vanish to all orders of τ. Thus if the leading term $A_0(x,t)$ in the amplitude is nonzero, then the phase function must satisfy the equation

$$\frac{1}{c(x)^2}\left(\frac{\partial\psi}{\partial t}\right)^2 - \sum_i \left(\frac{\partial\psi}{\partial x_i}\right)^2 = 0,$$

called the *eikonal* equation. The terms in the amplitude satisfy the recursive system of *transport* equations

$$2\frac{1}{c^2}\frac{\partial\psi}{\partial t}\frac{\partial A_k}{\partial t} - 2\sum_i \frac{\partial\psi}{\partial x_i}\frac{\partial A_k}{\partial x_i} + \frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2}A_k - \sum_i \frac{\partial^2\psi}{\partial x_i^2}A_k + q(A_{k-1}) = 0, \qquad k \geq 0.$$

These equations are solved consecutively for $k = 0, 1, \ldots$. For A_0 the term $q(A_{k-1})$ is absent. The transport equation describes the growth of the amplitude along the rays, the solution curves of the dynamical system

$$\frac{\partial x}{\partial\lambda} = \frac{\partial p}{\partial\xi}(x, \frac{\partial\phi}{\partial x}, \frac{\partial\phi}{\partial t}), \qquad \frac{\partial t}{\partial\lambda} = \frac{\partial p}{\partial\tau}(x, \frac{\partial\phi}{\partial x}, \frac{\partial\phi}{\partial t}),$$

where $p(x,\xi,\tau) = -\frac{1}{c(x)^2}\tau^2 + \|\xi\|^2$. This shows the connection with geometrical optics, since it allows for solutions localized around the rays.

The causal Green's function, which can be used to construct solutions of (1.1) for general f, can be approximated by an integral of solutions of the form (1.2). The causal Green's function $G(x, x_0, t)$ is the solution with f given by the Dirac δ-function $\delta(x - x_0)\delta(t)$, requiring that $G(x, x_0, t) = 0$ when $t < 0$. The solution for general f is given by

$$\int_0^t G(x, x_0, t - t_0) f(x_0, t_0) \, \mathrm{d}x_0 \, \mathrm{d}t_0.$$

We will approximate $G(x, x_0, t)$ by the following expression, involving an integral over the frequency $\tau \in \mathbb{R}$,

$$\int A(x, x_0, \tau) \mathrm{e}^{\mathrm{i}\tau(t - T(x,x_0))} \, \mathrm{d}\tau. \tag{1.3}$$

Here the phase function is given by $\tau(t - T(x, x_0))$. The eikonal equation for $T(x, x_0)$ reads in this case

$$c(x)^2 \sum_i \left(\frac{\partial T}{\partial x_i}\right)^2 = 1.$$

If we require as initial condition that $T(x, x_0) = \frac{1}{c(x)}\|x - x_0\| + O(\|x - x_0\|^2)$ when $x \to x_0$, then the solution is the traveltime from x_0 to x. This solution does not always exist. Let for the moment $x(x_0, \alpha, t)$ be the ray originating at x_0 with direction α at time t, then the solution exist as long as the map $S^{n-1} \times \mathbb{R}_+ \ni (\alpha, t) \mapsto x(x_0, \alpha, t)$ is a diffeomorphism. In other words caustics have to be absent.

When the differential operator P acts on the integrand in (1.3) this expression decays as τ^{-N} for any N. It follows that PG is smooth when $x \neq x_0$. There remains to chose m and the initial values for the A_k correctly, such that $PG = \delta(x - x_0)\delta(t)$ modulo some smooth function. Actually the smooth error term can also be removed by adding a smooth term to G, but we will not make use of this, because in the reconstruction problem the problems that occur with smooth terms are much harder and will not be treated here. By comparing with the constant coefficient case one finds $m = \frac{n-3}{2}$, and an expression for $A_0(x, x_0)$. All the A_k can be found by a different method where one writes the solution as an integral with multiple frequency variables such as given in (1.4), with $\theta = \xi$, the cotangent vector associated with x.

Because $m = 0$ for $n = 3$ we have the following expression for $G(x, x_0, t)$ in $n = 3$ dimensions

$$G(x, x_0, t) = (2\pi)A_0(x, x_0)\delta(t - T(x, x_0)) + \text{l.o.t.}$$

As already noted equation (1.3) is valid only when the wavefronts are smooth, and the rays from x_0 do not intersect each other. When this is not the case it turns out that we have to consider integrals of the following type

$$\int A(x,y,\theta)\mathrm{e}^{\mathrm{i}\phi(x,y,\theta)}\,\mathrm{d}\theta \tag{1.4}$$

where the frequency variable $\tau \in \mathbb{R}$ is replaced by a higher dimensional phase variable $\theta \in \mathbb{R}^N$. We require that the phase function ϕ is homogeneous of order 1 in θ, that is $\phi(x,\rho\theta) = \rho\phi(x,\theta)$, and that ϕ is nondegenerate. A phase function ϕ is called nondegenerate in a cone[2] $\Gamma \subset X \times \mathbb{R}^N$, where $X \subset \mathbb{R}^n$ contains the position variable x, if

$$\frac{\partial\phi}{\partial\theta}(x,\theta) = 0, \quad (x,\theta) \in \Gamma \qquad \Rightarrow \qquad \frac{\partial^2\phi}{\partial(x,\theta)\partial\theta}(x,\theta) \text{ has maximal rank.}$$

A way to see intuitively that (1.3) has to be replaced by an expression of the form (1.4) is by using Huygens principle. The wave front at some time t is given by an integral of contributions from each of the points of the wavefront at some earlier time t_0. The additional integration variables can be used as phase variables.

To understand in what sense certain distributions, such as the Green's function, can be approximated by expressions of the form (1.4) the concept of wave front set is required. We will introduce this following the presentation of Hörmander [25], Section 8.1. One can show that a distribution $v \in \mathcal{D}'(\mathbb{R}^n)$ is in $C_0^\infty(\mathbb{R}^n)$, if and only if its Fourier transform $\hat{v}(\xi)$ decays faster than polynomial when $|\xi| \to \infty$,

$$\hat{v}(\xi) \le C_N(1+|\xi|)^{-N}, \qquad N = 1,2,\dots,\ \xi \in \mathbb{R}^n. \tag{1.5}$$

Therefore there is a natural set of 'directions' associated to a distribution with singularities, consisting of those directions where the Fourier transform does not decay rapidly. Thus we define the cone $\Sigma(v)$ of all those $\eta \in \mathbb{R}^n\backslash 0$ having no conic neighborhood V such that (1.5) is valid when $\xi \in V$. It is clear that $\Sigma(v)$ is a closed cone in $\mathbb{R}^n\backslash 0$, and that $\Sigma(v) = \emptyset$ if and only if $v \in C_0^\infty$. It is shown in Lemma 8.1.1 of Hörmander [25] that if $\phi \in C_0^\infty(\mathbb{R}^n)$, then $\Sigma(\phi v) \subset \Sigma(v)$. It follows that we can define for $u \in \mathcal{D}'(X)$

$$\Sigma_x(u) = \bigcap_\phi \Sigma(\phi u), \qquad \phi \in C_0^\infty(X), \qquad \phi(x) \neq 0.$$

The wavefront set of a distribution is now defined as

$$\mathrm{WF}(u) = \{(x,\xi) \in X \times (\mathbb{R}^n\backslash 0) \,|\, \xi \in \Sigma_x(u)\}.$$

Thus singularities of a distribution $u \in \mathcal{D}'(X)$ can be localized not only in the space X (using the singular support), but even in the cotangent space $T^*X\backslash 0$. This is called microlocalization. Let $\Gamma \subset T^*X\backslash 0$ be an open cone. We say $u = v$ microlocally

[2]By conic set or cone Γ we mean that $(x,\theta) \in \Gamma$, implies that $(x,\rho\theta) \in \Gamma$ for all $\rho > 0$.

in Γ, if $\mathrm{WF}(u-v) \cap \Gamma = \emptyset$. Such microlocal equivalence is always modulo smooth functions.

Typically an expression such as (1.4), called a local Fourier integral operator, describes an operator microlocally. The whole operator can be given by a sum of local Fourier integral operators, modulo an operator with smooth kernel. For example, using integral operators with a kernel of the form (1.4), distributions can be given that are equivalent to the solution of a wave equation microlocally in some conic subset of $X \times \mathbb{R}^n \backslash 0$.

The theory of Fourier integral operators describes the properties of operators with a kernel of the form (1.4). An important feature is that the choice of phase function and amplitude is in general far from unique. Using the method of stationary phase the number of phase variables can sometimes be reduced, and the phase function and amplitude can also be changed by applying a coordinate transformation to the phase variable.

To deal with this nonuniqueness, invariant objects are defined. The invariant quantity related to the phase function of the operator G is the canonical relation, that describes the relation between the wave front sets of a distribution f and the distribution Gf. It is a subset of $T^*X\backslash 0 \times T^*Y\backslash 0$. An invariant characterization of the highest order part of the amplitude $A(x, y, \theta)$ is given by the principal symbol.

Using these invariant quantities a calculus can be defined. It shows that under certain conditions on the canonical relations the composition of two Fourier integral operators is again a Fourier integral operator, such that the canonical relation of the composition is simply the composition of the canonical relations.

For further information about Fourier integral operators we refer to the books by Duistermaat [17], Hörmander [25, 26, 27], Treves [57, 58] and Maslov and Fedoriuk [38].

1.2 The seismic inversion problem

In this section we discuss some important ideas in the inversion of seismic data. We discuss the modeling and inversion of seismic data in simple acoustic media. It is based for a large part on the paper by Beylkin [8].

Before the inversion takes place usually some preprocessing is applied, to remove unwanted features of the data. So we will assume that the source wavelet is approximately a δ-function. We assume that the influence of the near surface, that varies strongly from location to location, is removed. We will also assume that the effects of the boundary in the data are suppressed. Thus surface waves are removed. This can be done by a cutoff in the Fourier domain. Also data related to waves that have reflected at the surface are removed. This can be done approximately without knowledge of the medium coefficients. So we assume that in effect the sources and receivers are located on plane in a medium that extends on both sides of the plane.

Assuming that the discontinuities in the medium coefficients are small, the following linearization in the medium coefficients can be done. The full soundspeed is written as a sum

$$c(x) + \delta c(x), \tag{1.6}$$

where the background medium $c(x)$ is smooth and the perturbation $\delta c(x)$ contains the singularities. Here and in the sequel we use the classic notation of $\delta c, \delta P, \delta G, \ldots$ for 'small perturbation of $c, P, G, \ldots$'. The resulting perturbation of the Green's function is given by

$$\delta G(r, s, t) = -\int G(r, x, t - t_0)\, \delta P\, G(x, s, t_0)\, \mathrm{d}x_0\, \mathrm{d}t_0. \tag{1.7}$$

Here s is the source coordinate, r the receiver coordinate, and t the time. The second approximation consists of the calculation of the Green's function in the high-frequency approximation. Then (1.7) describes singly reflected waves. We will assume that direct rays from source to receiver are absent, so that (1.7) actually describes all the data.

These two approximations are often done together and are called ray+Born approximation. The approximation implies that one assumes separation of scales in the medium, in the sense that the length scale of the variations in the smooth part should be much larger than the length scale of the variations in the medium perturbation. Indeed the wavelength of the reflected signal is of the order of the length scale of the medium perturbation, while the variations of the smooth part are assumed to be on a larger scale.

Although the ray+Born approximation gives an efficient way to model scattering data, it is not quite satisfactory. For instance, it is not clear how accurate the ray+Born approximation is. Moreover, measurements of real media in boreholes indicate that the assumption of separation of scales need not be satisfied. Propagation of waves in inhomogeneous media is still an active area of research.

Using expression (1.3) in the equation for the perturbation of the Green's function (1.7) we obtain

$$\delta G(r, s, t) = \int A(x, s, \tilde{\tau}) A(x, r, \hat{\tau}) \tilde{\tau}^2 \mathrm{e}^{\mathrm{i}\hat{\tau}(t - t_0 - T(x,r)) + \mathrm{i}\tilde{\tau}(t_0 - T(x,s))} \delta(\frac{1}{c(x)^2})\, \mathrm{d}\hat{\tau}\, \mathrm{d}\tilde{\tau}\, \mathrm{d}t_0\, \mathrm{d}x. \tag{1.8}$$

Define the linearized forward operator F to be the operator that maps the medium perturbation $\delta\frac{1}{c(x)^2} = -2c(x)^{-3}\delta c(x)$ (here we take the coefficient in δP, instead of δc itself) to $\delta G(r, s, t)$

$$F : -2c(x)^{-3}\delta c(x) \mapsto \delta G(r, s, t)$$

given by (1.8).

We assume that $\delta c = 0$ on a neighborhood of the set of source and receiver positions. Then a cutoff function can be introduced to smoothly cut off the integral at $t_0 = 0$ and $t_0 = t$. The right-hand side of (3.37) can now be simplified by doing the integral over t_0 and one of the τ variables, using the method of stationary phase. We explain this in detail. The phase in the integral can be rewritten as

$$(t - T(x,r) - T(x,s))\frac{\hat{\tau} + \tilde{\tau}}{2} + (t_0 + \tfrac{1}{2}(T(x,r) - T(x,s) - t))(\tilde{\tau} - \hat{\tau}).$$

Define new variables $\tau = \frac{\hat{\tau}+\tilde{\tau}}{2}, \sigma = \tilde{\tau} - \hat{\tau}$. We write the right-hand side of (1.8) as

$$\delta G(r,s,t) = \int B(x,s,r,\tau) \mathrm{e}^{\mathrm{i}(t-T(x,r)-T(x,s))\tau} \, \mathrm{d}\tau \, \mathrm{d}x,$$

where we have defined a new amplitude

$$B(x,s,r,\tau) = \int A(x,s,\tilde{\tau}(\tau,\sigma)) A(x,r,\hat{\tau}(\tau,\sigma)) \tilde{\tau}(\tau,\sigma)^2 \mathrm{e}^{\mathrm{i}(t_0 + \frac{1}{2}(T(x,r)-T(x,s)-t))\sigma} \, \mathrm{d}\sigma \, \mathrm{d}t_0.$$

The singularities of $\delta G(r,s,t)$ depend only on the the asymptotic behavior when $\rho \to \infty$ of $B(x,s,r,\rho\tau)$. This expression reads (changing variables from σ to $\rho\sigma$)

$$\begin{aligned} B(x,s,r,\tau) = &\int A(x,s,\rho\tilde{\tau}(\tau,\sigma) A(x,r,\rho\hat{\tau}(\tau,\sigma))(\rho\tilde{\tau}(\tau,\sigma))^2 \\ &\times \mathrm{e}^{\mathrm{i}\rho(t_0+\frac{1}{2}(T(x,r)-T(x,s)-t))\sigma} \, \rho \, \mathrm{d}\sigma \, \mathrm{d}t_0. \end{aligned}$$

We can now apply the stationary phase formula, see e.g. Duistermaat [17], Proposition 1.2.4, with $t = \rho, x = (t_0, \sigma)$. The end result is

$$\delta G(r,s,t) = 2\pi \int A(x,s,\tau) A(x,r,\tau) \tau^2 \mathrm{e}^{\mathrm{i}\tau(t-T(x,r)-T(x,s))} \delta(\frac{1}{c(x)^2}) \, \mathrm{d}\tau \, \mathrm{d}x. \tag{1.9}$$

One can check that under the condition that incoming and outgoing rays intersect transversally (i.e. there is no direct ray from source to receiver), the phase function in (1.9) is nondegenerate and the map $-2c(x)^{-3}\delta c(x) \mapsto \delta G(r,s,t)$ is a Fourier integral operator.

In $n = 3$ dimensions this reads

$$\delta G(r,s,t) = -2\pi \int A_0(x,s) A_0(x,r) \delta''(t - T(x,r) - T(x,s)) 2c(x)^{-3} \delta c(x) \, \mathrm{d}x + \text{l.o.t.}$$

Thus the map from $\delta c(x)$ to the data can be described as follows. One takes the integral of a second derivative of the data multiplied with some amplitude over a set of x where the total traveltime is constant $T(x,r) + T(x,s) = t$, called the isochron. Because of this integration over surfaces this transformation is sometimes called a generalized Radon transform (GRT)[3].

[3]We recall the definition of the classical Radon transform. If $\theta \in S^{n-1}, p \in \mathbb{R}$, then the classical Radon transform maps a function f of $x \in \mathbb{R}^n$ to its integrals over hyperplanes in $\mathbb{R}^n$, $Rf(\theta,p) = \int_{\langle x,\theta \rangle = p} f(x) \, \mathrm{d}x$.

Recall that the data are modeled by the perturbation of the Green's function $\delta G(r, s, t)$. The formulas above show that the data are modeled by a linear operator F acting on the medium perturbation $\delta c(x)$, where F depends on the background medium. Therefore, let us first try to invert for $\delta c(x)$. The first thing that should be noticed, is that there is a redundancy in the data. The data are described by a function of $2n-1$ variables, while the function $\delta c(x)$ is a function of n variables. Consider therefore subsets of data with $n-1$ variables fixed. It is common to fix what is called the offset coordinate $h = r - s$. We define F_h to be the operator mapping $\delta c(x) \mapsto \delta G(s+h, s, t)$. Let

$$S_h(s, x, \tau) = -\tau(T(x, s+h) + T(x, s)),$$

the phase function minus the factor $t\tau$.

It has been shown by Beylkin [7, 8] that inversion for $\delta c(x)$ modulo smooth terms is possible if the following equation is satisfied

$$\det \frac{\partial^2 S_h}{\partial x \partial(s, \tau)} \neq 0. \tag{1.10}$$

In that case the map $(s, \tau) \mapsto \frac{\partial S_h}{\partial x}$ is locally a submersion. We assume that the condition is satisfied for some open set of x and some open set of s. In that case at least microlocally (see Section 1.1) an inverse is obtained.

Theorem 1.2.1 *(Beylkin) If (1.10) is valid on some open set of (s, x), then F_h is microlocally invertible around $(x, \frac{\partial S}{\partial x})$.*

Proof The idea of the proof is to show that $F_h^* F_h$ is a pseudodifferential operator with invertible symbol. Then an asymptotic inverse is given by

$$F_h^{-1} = (F_h^* F_h)^{-1} F_h^*.$$

For the details we refer to Beylkin [7, 8]. □

If this inverse exists for each value of the offset h we obtain in fact an invertible map that maps seismic data to a function $r(x, h) = -2c(x)^{-3}\delta c(x)$. The reconstructed medium perturbation should of course not depend on the offset h that is used. This will be the case if the background medium is correctly chosen. However, if the background medium is incorrect there will be some dependence on h. In this way the redundancy in the data can be used to reconstruct the background medium. We formulate the result as a theorem.

Theorem 1.2.2 *The perturbation of the Green's function $\delta G(r, s, t)$ that models the data, is given by an operator H acting on a function $r(x, h)$*

$$H : r(x, h) \mapsto d(r, s, t).$$

Here $r(x, h)$ is such that the position of its singularities does not depend on h. The operator H can be constructed as a Fourier integral operator that is invertible microlocally. For the Born approximation in acoustic media $r(x, h) = -2c(x)^{-3}\delta c(x)$.

To test whether a function $r(x,h)$, obtained from inverting data, does not depend on h one can take the derivative $\frac{\partial r}{\partial h}(x,h)$. This derivative should be a smooth function (recall that the inversion is only modulo a smooth function). Conjugating the operator $\frac{\partial}{\partial h}$ with H we obtain an operator on data, that should result in a smooth function. We formulate this as a corollary.

Corollary 1.2.3 *The operator*

$$H\frac{\partial}{\partial h}H^{-1}$$

is a pseudodifferential operator which microlocally annihilates the data.

A square integral of this operator acting on data may be used as a measure for the error in the choice of the background medium. The question now arises whether by minimization of this error the background medium can be found. This procedure is called differential semblance optimization. It has been discussed by Symes [55]. It leads to a nonlinear inverse problem that is very different from the reconstruction problem for the singular part of the medium. This problem will not be discussed in this thesis.

1.3 Contents of this thesis

1.3.1 Wave equations with discontinuous coefficients

As was explained above, the subsurface can be modeled as an acoustic or elastic medium, with coefficients that may have discontinuities. In fact measurements of the local wavespeed in wells indicate that the coefficients are far from smooth and vary on many scales. Since most textbooks discuss the acoustic and elastic equations only with smooth coefficients we wondered whether in these cases solutions could still be defined.

In Chapter 2 we show, following Lions and Magenes [36], that the initial value problem for wave equations is well-posed when the coefficients are only measurable and bounded away from zero and infinity. We show that when initial values and source belong to suitable function spaces, such as the Sobolev space $H_{(1)}$ of order one, then solutions exist that are continuous w.r.t. time with values in $H_{(1)}$ (for the space coordinates). To obtain these results the second order partial differential equation is written as an ordinary equation with values in a Hilbert space. Energy estimates are given that lead to the mentioned results.

There is a scale of spaces, such that solutions can be found that take their values in these spaces. These spaces are defined as the domain of the spatial part of the differential operator and its powers and as the spaces dual to these domains. Here the coefficient in front of the time derivative is normalized to one. An interesting feature of the dual spaces is that these are not distribution spaces in the ordinary sense. For coefficients that are discontinuous on a plane we give an explicit form of these spaces, and we relate it to the more conventional formulation involving conditions at the interface. As far as we know this result is new.

The last two sections are about the dependence of the solutions on the coefficients. It is known that under certain conditions the solutions depend continuously on the coefficients, when the perturbation of the latter is measured in a suitable way. This is used to obtain differentiability of the dependence of the solutions on the coefficients. We have not seen this last result in the literature. It is of importance to understand the linearization done in (1.6), (1.7). An interesting feature of the result is that the solution to a wave equation modeled by linearization converges to the real solution in the L^2 norm, and not in the norm of the space $H_{(1)}$ that contains the solutions. This suggests (but this should be investigated in more detail) that when using an iterative procedure to determine the medium parameter $c(x)$ one should minimize the difference

$$\text{real data} \ - \ \text{approximated data}$$

in a norm that is different from the norm of the space that contains the data, unlike what one might think naievely.

1.3.2 Modeling and inversion of seismic data in anisotropic elastic media

Chapter 3 contains joint work with M.V. de Hoop [53]. We give a generalization of Theorem 1.2.2 and Corollary 1.2.3 in two directions: we consider media with caustics, and we consider elastic media. We discuss modeling using both the Born approximation introduced above, and the Kirchhoff approximation, where the medium coefficients are assumed to be piecewise smooth and only have jump discontinuities along smooth interfaces. We give a high-frequency analysis using Fourier integral operators. Maximal datasets (sources and receivers on an open part of the boundary) as well as nonmaximal datasets are treated. We discuss general (isotropic or anisotropic) media.

The result Theorem 3.7.1 is formulated for elastic media, but because of the similarities with the acoustic case we first describe the situation for acoustic media. For acoustic media with caustics there is already a substantial literature, mostly discussing least-squares inversion for the Born approximation. Rakesh [46] shows that the linearized forward operator is a Fourier integral operator, Hansen [22] and Ten Kroode, Smit and Verdel [34], discuss the least-squares inversion (the latter discuss in detail the conditions that are involved). Nolan and Symes [41] consider the case where the set of source and receiver positions is not maximal. The reconstruction of a jump discontinuity at an interface, given the medium above interface, is given by Hansen [22].

The disadvantage of the least-squares approach is that the redundancy in the data is not characterized, and is not used fully. In acoustic media the redundancy is present because reflection occurs for a range of values of the scattering angle (the angle between incoming and outgoing rays at the scattering point). Therefore the idea is to use the scattering angle to describe the redundancy. In Chapter 3 we generalize Theorem 1.2.2 in the sense that we construct an operator H microlocally that maps the data to a function $r(x, e)$, where in acoustic media e can always be chosen as the scattering angle. The function $r(x, e)$ is such that the singular part of the function essentially does not depend on e. The operator H is invertible. Although the idea of transforming to the scattering angle domain existed before, such a transformation was in the presence of caustics only defined for a subset of the data.

The propagation of waves in elastic media is described by a system of partial differential equations. In a constant coefficient medium one can show that there are longitudinal and transversal waves that propagate independently, i.e. the system decouples. In media with smoothly varying coefficients this decoupling is no longer exact, but in the high-frequency limit it is in most cases still valid (at least microlocally, see Taylor [56], Ivrii [30], Dencker [16], this is discussed in more detail in Section 3.2.1).

In Chapter 3 we first discuss propagation of waves in the Kirchhoff approximation, i.e. in media with jump discontinuities along smooth interfaces. We then discuss the Born approximation and give a least-squares approach. We give a new presentation, different from the one in the paper by De Hoop and Brandsberg-Dahl [24]. In order to generalize Theorem 1.2.2 (both in the acoustic and the elastic case), the canonical relation of the modeling operators is analyzed. The canonical relation of the operator H is constructed, as well as phase functions related to phase functions of the original operators. Combining the different results then leads to our main result, which is the equivalent of Theorem 1.2.2 for elastic media, given in Theorem 3.7.1.

It turns out that essentially one can reconstruct, for each pair of elastic modes (M, N), a reflectivity function $r_{MN}(x, e)$ that is a function of subsurface position and an additional coordinate describing the redundancy (in many cases, but not always this can be the scattering angle). Thus the position of a reflector and the reflection coefficient as a function of direction of incoming and outgoing wave can be reconstructed. The equivalent of Corollary 1.2.3 is given in Corollary 3.7.3.

1.3.3 Linearized inversion when traveltime injectivity is violated

Chapter 4 contains the work [52]. It discusses inversion of the high-frequency medium perturbation of an acoustic medium in the Born approximation (the function $\delta c(x)$ in (1.6)). It is shown by Hansen [22] and by Ten Kroode, Smit and Verdel [34] that under the so called traveltime injectivity condition a least-squares inverse for $\delta c(x)$ exists. This condition is a geometrical condition on the rays, explained in detail in the article by Ten Kroode, Smit, Verdel [34] and also in Section 4.3 of this thesis. In our work we discuss the situation where the condition is violated.

In the Born approximation the data can be modeled by a Fourier integral operator F acting on the medium perturbation $\delta c(x)$. To find a least-squares inverse the normal operator $N = F^*F$ is considered. An inverse for F exists if and only if N is invertible and in that case the least-squares inverse is given by $F^{-1} = N^{-1}F^*$. It has been shown that if the traveltime injectivity condition is satisfied, then N is an invertible pseudodifferential operator of order $n-1$, where n is the dimension of the subsurface. If the traveltime injectivity condition is violated, then the normal operator is the sum of such a pseudodifferential operator and a nonmicrolocal part. Ten Kroode et al. [34] show that under a certain condition this nonmicrolocal part is a Fourier integral operator of order $\frac{n-1}{2}$.

Our first result is that we simplify this condition. Also we show that it is satisfied generically. Next we note that although the nonmicrolocal part is of lower order than the pseudodifferential part, this does not imply that it is also more regular as an operator between Sobolev spaces. It follows that Theorem 3.2 in Ten Kroode et al. is too optimistic concerning the regularity of the nonmicrolocal part. Using known estimates for so called degenerate Fourier integral operators, we analyze the regularity of the nonmicrolocal part between Sobolev spaces in detail. We construct examples where the nonmicrolocal part is as singular as the pseudodifferential part, and the inversion is, at least microlocally, no longer possible. Finally we show that generically the nonmicrolocal part is more regular than the pseudodifferential part, so that generically the linearized forward operator still has an asymptotic inverse.

Chapter 2

Wave equations with discontinuous coefficients

2.1 Introduction

In this chapter we study the propagation of waves in media with discontinuous coefficients. Examples of the systems that we study are the acoustic and elastic equations. The reason to devote a chapter to this topic is that most texts on wave equations discuss only equations with smooth coefficients. In this chapter we present a number of results that have appeared in the mathematical literature and also some results that we have not seen elsewhere, and that may be new (mostly in Sections 2.7 and 2.9).

When the coefficients are discontinuous the partial differential equation has to be interpreted in some weak sense, i.e. by integrating with a test function, to which some of the differentiations are transposed using partial integration. In Section 2.2 it is shown that this can be done when the coefficients are in L^∞ (measurable and bounded) and the solutions are in the Sobolev space $H_{(1)}$, which consists of the L^2 functions with derivatives that are also in L^2. Moreover, the partial differential equation is then treated as an ordinary differential equation for a function of the time variable with values in a Hilbert space of functions of the spatial variables. In Section 2.3 this is done for domains with a boundary.

If in addition the coefficients are bounded away from zero, then the theory of Lions and Magenes [36], Section III.8, can be used to show that the initial value problem is well posed. This is done in Section 2.4. Using the method of energy estimates we show that there is existence, uniqueness and stability of solutions and that the solutions are continuous functions of t with values in the Sobolev space $H_{(1)}$. In Section 2.5 a similar result is shown for more general equations with some weak kind of time dependence and lower order terms.

In Section 2.6 solutions are discussed with more or less regularity than the solutions given in Section 2.4. For equations with smooth coefficients it is natural to work in the class of Sobolev spaces $H_{(s)}$. The regularity of solutions to wave equations depends on the regularity of initial values and of the source. For equations with nonsmooth coefficients these Sobolev spaces are replaced by spaces that can roughly be described as the domain of powers of the space part of the wave operator.

In some sense the simplest way that the coefficients can be discontinuous is when they have a jump discontinuity at a codimension one hypersurface. This case is discussed in Section 2.7. It is known that the solutions to such equations should satisfy certain continuity conditions at the discontinuity. We show how these conditions are related to the distributional form of the differential equation. This theory is used to model seismic data in Chapter 3.

In the reconstruction problem, and for the approximation of solutions, it is important to know how the solutions depend on the coefficients. In Section 2.8 we show that the solutions (given the initial values and the source) depend continuously on the coefficients, when the latter converge in measure. In Section 2.9 we show that the solutions depend differentiably on the coefficients in L^∞. So the solutions can be approximated in some norm by linearization of the coefficients. To model seismic data often such a linearization is used, see for instance Chapters 3 and 4.

To obtain our results we use the method of energy estimates. Some functional analysis and integration theory is used. Useful references include Lions and Magenes [36], Dunford and Schwartz [18], Bourbaki [11], and Hörmander [25].

Some of the results may also be obtained using the theory of semigroups (see for instance Yosida [61] or Pazy [42]), or their second order version, cosine theory (see Fattorini [20]). The continuous dependence on the coefficients then follows from the Trotter-Kato theorem (see Yosida [61], Section 9.12). Another way to obtain existence and uniqueness results is by regularizing the coefficients and using energy estimates and an existence result for equations with smooth coefficients. One can use the theory of Section 2.8 to define the solutions to an equation with coefficients in L^∞ as a limit of solutions to equations with smooth coefficients. This is done by Hurd and Sattinger [28].

2.2 The partial differential equation

In this section we discuss partial differential equations such as the acoustic and elastic wave equations with very weak assumptions on the coefficients. We only require that they are in L^∞ and bounded away from zero. This will require the use of weak solutions. We then show how these equations can be seen as an ordinary differential equation for a function that takes values in a Hilbert space. This will allow us to use the theory of later sections which shows that the initial value problem is well posed in the sense of Hadamard.

We will mostly be interested in the following system of equations for a function $u_K(x,t)$ that takes its values in $\mathbb{C}^N$

$$-\sum_{i,j,L} \tfrac{\partial}{\partial x_i} a_{ij;KL}(x) \tfrac{\partial}{\partial x_j} u_L(x,t) + \sum_L \tfrac{\partial}{\partial t} c_{KL}(x) \tfrac{\partial}{\partial t} u_L(x,t) = f_K(x,t). \tag{2.1}$$

Here i, j denote the component of the spatial coordinate x, K, L indicate the component of the vector u_K, and $(x,t) \in \mathbb{R}^n \times]0,T[$ (since the wave equation is symmetric under time reversal $t \to -t$ we could also choose $t \in]-T,T[$). The coefficients $a_{ij;KL}(x), c_{KL}(x)$ are selfadjoint for $i \leftrightarrow j$ and $K \leftrightarrow L$, $a_{ij;KL} = \overline{a_{ji;LK}}$. We also assume they are positive definite in the sense that there are constants $\alpha, \beta > 0$ such that for all $x \in \mathbb{R}^n$ and for all ξ_i, u_K

$$\sum_{i,j,K,L} a_{ij;KL} \overline{\xi_i} \xi_j \overline{u_K} u_L \geq \alpha \|\xi\|^2 \|u\|^2,$$
$$\sum_{K,L} c_{KL} \overline{u_K} u_L \geq \beta \|u\|^2. \tag{2.2}$$

The acoustic and elastic equations are special cases of this. For the acoustic equation $N = 1$, u is the pressure, $a_{ij}(x)$ is given by $\rho(x)^{-1}\delta_{ij}$ ($\rho(x)$ the mass density) and $c(x)$ is given by $\kappa(x)$, the compressibility. For the elastic system $N = n$, $u_K(x,t)$ is the displacement vector, $a_{ij;KL}(x)$ is given by the elastic tensor $c_{iKjL}(x)$, and $c_{KL}(x)$ is given by $\rho(x)\delta_{KL}$, where again $\rho(x)$ is the mass density.

Since we are especially interested in the situation where the coefficients vary discontinuously, this equation should be interpreted distributionally. Hence equation (2.1) will be understood in the weak sense that

$$\int \left(- \sum_{i,j,K,L} \overline{v_K(x,t)} \tfrac{\partial}{\partial x_i} a_{ij;KL}(x) \tfrac{\partial}{\partial x_j} u_L(x,t) + \sum_{K,L} \overline{v_K(x,t)} \tfrac{\partial}{\partial t} c_{KL}(x) \tfrac{\partial}{\partial t} u_L(x,t) \right) \mathrm{d}x \, \mathrm{d}t$$
$$= \int \sum_K f_K(x,t) \overline{v_K(x,t)} \, \mathrm{d}x \, \mathrm{d}t, \tag{2.3}$$

for all test functions $v_K \in C_0^\infty(\mathbb{R}^n \times [0,T], \mathbb{C}^N)$. By partial integration we find that, in the case that u_L and $a_{ij;KL}$ are smooth, this is equivalent to

$$\int \left(\sum_{i,j,K,L} a_{ij;KL}(x) \frac{\partial u_L}{\partial x_j}(x,t) \frac{\partial \overline{v_K}}{\partial x_i}(x,t) - \sum_{K,L} c_{KL}(x) \frac{\partial u_L}{\partial t}(x,t) \frac{\partial \overline{v_K}}{\partial t}(x,t) \right) \mathrm{d}x \, \mathrm{d}t$$
$$= \int \sum_K f_K(x,t) \overline{v_K}(x,t) \, \mathrm{d}x \, \mathrm{d}t, \tag{2.4}$$

for all test functions $v_K \in C_0^\infty(\mathbb{R}^n \times [0,T], \mathbb{C}^N)$. The point is that (2.4) also makes sense for nonsmooth u_L and $a_{ij;KL}$, in particular for $u_L \in \bar{H}_{(1)}(\mathbb{R}^n \times]0,T[)$, $a_{ij;KL}$ bounded and measurable, and $f_K \in L^2(\mathbb{R}^n \times]0,T[)$. For this reason we use (2.4) instead of (2.1) as our equation of motion. Here the Sobolev space $\bar{H}_{(m)}(\Omega)$ is defined for integer order $m \geq 0$, and for a domain with a boundary Ω, as

$$H_{(m)}(\Omega) = \{ u | \partial^\alpha u \in L^2(\Omega) \ \forall \alpha, |\alpha| \leq m \}, \tag{2.5}$$

where the derivative is taken in the sense of distributions (Sobolev spaces on domains with a boundary are discussed in more detail in Section 2.3). It is a Hilbert space for the inner product

$$\langle u, v\rangle = \int \sum_{|\alpha|\leq m} \frac{\partial^\alpha u}{\partial x^\alpha}(x)\frac{\partial^\alpha \overline{v}}{\partial x^\alpha}(x)\,\mathrm{d}x,$$

where α is a multi-index $x^\alpha = (x_1^{\alpha_1}, \dots, x_n^{\alpha_n})$.

Remark 2.2.1 In fact the acoustic and elastic equations can be derived in the form (2.4), by using Lagrange's variational formulation of classical mechanics.

We are interested in the Cauchy problem, we want to find $u_K(x,t)$ that satisfy (2.4) and the initial conditions

$$u_K(x,0) = u_{0,K}(x), \qquad\qquad \frac{\partial u_K}{\partial t}(x,0) = u_{1,K}(x),$$

with $u_{0,K} \in H_{(1)}(\mathbb{R}^n, \mathbb{C}^N), u_{1,K} \in L^2(\mathbb{R}^n, \mathbb{C}^N)$. Below equation (2.19) we argue that $u_K(x,0)$, $\frac{\partial u_K}{\partial t}(x,0)$ are indeed well defined.

It turns out to be very useful to view $u_K(x,t)$ as a function $t \mapsto u(t)$ with values $u(t) : x \mapsto u_K(x,t)$ in a Hilbert space of functions in $\mathbb{R}^n$. The Fubini theorem says that

$$L^2(X \times Y) \sim L^2(X, L^2(Y)).$$

From this we get, for a measurable function u on $\mathbb{R}^n \times]0,T[$

$$u \in \bar{H}_{(1)}(\mathbb{R}^n \times]0,T[, \mathbb{C}^N) \quad \Leftrightarrow \quad \begin{cases} u \in L^2(]0,T[, H_{(1)}(\mathbb{R}^n, \mathbb{C}^N)), \\ u' \in L^2(]0,T[, L^2(\mathbb{R}^n, \mathbb{C}^N)). \end{cases}$$

We define the spaces $V = H_{(1)}(\mathbb{R}^n, \mathbb{C}^N)$ and $H = L^2(\mathbb{R}^n, \mathbb{C}^N)$.

Let the operators A, C be defined by

$$A = -\sum_{i,j} \tfrac{\partial}{\partial x_i} a_{ij;KL}(x) \tfrac{\partial}{\partial x_j}, \qquad\qquad C = c_{KL}(x). \tag{2.6}$$

As we see in (2.4) there are symmetric sesquilinear forms on V and H, associated with A and C respectively. We define

$$\begin{aligned} a(u,v) &= \int \sum_{i,j,K,L} a_{ij;KL}(x) \frac{\partial u_K}{\partial x_i}(x) \frac{\partial \overline{v_L}}{\partial x_j}(x)\,\mathrm{d}x, \qquad u, v \in V, \\ c(u,v) &= \int \sum_{K,L} c_{KL}(x) u_K(x) \overline{v_L}(x)\,\mathrm{d}x, \qquad u, v \in H. \end{aligned} \tag{2.7}$$

We have

$$\langle Au, v\rangle = a(u, v) = \langle u, Av\rangle,$$

and it follows that $A : V \to V'$, where V' is the dual of V. Also

$$\langle Cu, v\rangle = c(u, v) = \langle u, Cv\rangle.$$

With these definitions the differential equation in variational form reads

$$\int_0^T [a(u(t), v(t)) - c(u'(t), v'(t))]\,\mathrm{d}t = \int_0^T \langle f(t), v(t)\rangle\,\mathrm{d}t. \tag{2.8}$$

for all $v \in C_0^\infty([0, T], V)$.

From the estimates (2.2) we obtain that

$$c(u, u) \geq \beta\|u\|_H^2, \tag{2.9}$$

and, for $\lambda > \alpha$,

$$a(u, u) + \lambda\|u\|_H^2 \geq \alpha\|u\|_V^2. \tag{2.10}$$

Often it will be implicit that equations are in a distributional sense. For the PDE we write in that case simply

$$Au(t) + C\frac{\partial^2 u}{\partial t^2}(t) = f(t).$$

2.3 The PDE on domains with a boundary

If the domain X has a boundary ∂X then a boundary condition has to be imposed. Common conditions are the Dirichlet or Neumann conditions, given by

$$u_K(x, t) = 0 \qquad \text{on } \partial X, \tag{2.11}$$

or

$$\sum_{i,j,L} \nu_i a_{ij;KL} \frac{\partial u_L}{\partial x_j} = 0 \qquad \text{on } \partial X, \tag{2.12}$$

respectively, where ν_i is the normal to the boundary. Here it is assumed that u is in a class of distributions such that the restriction to ∂X can be performed. We write the problem with conditions on a smooth boundary in the form (2.8), where a, c are again sesquilinear forms satisfying (2.10), (2.9). Here we have to choose suitable Sobolev spaces V, H. Sobolev spaces can be defined on arbitrary domains in $\mathbb{R}^n$, see Adams [1], and Lions and Magenes [36]. It follows that the weak form can also be given for arbitrary domains in $\mathbb{R}^n$.

Spaces of distributions on the domain $X = \mathbb{R}^n_+$ with boundary $\partial X = \mathbb{R}^{n-1}$ have been discussed by Hörmander [26], appendix B.2. If $F(\mathbb{R}^n)$ is a space of distributions then $\bar{F}(X)$ is defined as the space of restrictions to X of its elements, and $\dot{F}(\overline{X})$ as the space of distributions supported by $\overline{X}$. For example one can show that $\bar{C}_0^\infty(X)$ consists of the elements of $C^\infty(\overline{X})$ that vanish outside a compact set. These functions can be extended smoothly across the boundary. The space $\dot{C}_0^\infty(\overline{X})$ is the subspace of functions vanishing of infinite order at ∂X. For integer $k \geq 0$ the space $\bar{H}_{(k)}(X)$ may be defined as in (2.5). The space $\bar{C}_0^\infty(\mathbb{R}^n_+)$ is dense in $\bar{H}_{(k)}(X)$. The space $\dot{H}_{(k)}(X)$ may be defined as the closure of $C_0^\infty(X)$ in $\bar{H}_{(k)}(X)$. The spaces $\dot{H}_{(-k)}, \bar{H}_{(-k)}$ are the duals of $\bar{H}_{(k)}, \dot{H}_{(k)}$. These definitions can be extended to the case where X is an open subset of $\mathbb{R}^n$ with smooth boundary and also to the case where X is a manifold with boundary.

In the Dirichlet case we define

$$V = \dot{H}_{(1)}(X), \qquad\qquad H = L^2(X).$$

The sesquilinear forms a, c are defined as in (2.7). Our set of test functions in this case is $C_0^\infty([0,T], V)$, so they vanish on the boundary. It follows by integration by parts that the original partial differential equation in weak sense (2.3) is equivalent to (2.8).

For this case the estimates (2.9), (2.10) are satisfied. In fact if X is bounded the constant λ can be chosen to be 0. This can be seen most directly in the one dimensional case. We show that in this case

$$\|u\|_{L^2} \leq C\|u'\|_{L^2}. \tag{2.13}$$

Suppose $u(x)$ has support in $[0, s]$, for some s. Clearly

$$|u(x)| \leq \int_0^x |u'(y)| \,\mathrm{d}x$$

The right hand side can be seen as the L^2 inner product between the indicator function $1_{[0,x]}$ and the function $|u'|$, so we have

$$u(x) \leq \|1_{[0,x]}\|_{L^2}\|u'\|_{L^2} = \sqrt{x}\|u'\|_{L^2} \leq \sqrt{s}\|u'\|_{L^2},$$

so $u(x)$ is bounded and since the domain is bounded (2.13) follows. This argument can be extended to the multidimensional case.

For the Neumann boundary condition (2.12) we define

$$V = \bar{H}_{(1)}(X), \qquad\qquad H = L^2(X),$$

and the sesquilinear forms a, c are unchanged. In this case it is not clear a priori that the restriction of $\sum_{i,j,L} \nu_i a_{ij;KL} \frac{\partial u_L}{\partial x_j}$ to the boundary can be taken. This will be shown first. After that we argue that the partial differential equation with the Neumann boundary condition is equivalent to (2.8).

Using a C^1 coordinate transformation we may assume that X is locally given by $\mathbb{R}^n_+$. Suppose $u \in \bar{H}_{(1)}(X\times]0,T[)$ satisfies the PDE in weak form with some $f \in L^2(X\times]0,T[)$. We show that $\sum_{j,L} a_{nj;KL}\frac{\partial u_L}{\partial x_j}$ can be seen as a continuous function of x_n with values in a space of distributions, so that the restriction to $x_n = 0$ is a well defined distribution. First note that

$$\sum_{j,L} a_{nj;KL}\frac{\partial u_L}{\partial x_j} \in L^2(\mathbb{R}^n_+\times]0,T[) = L^2(\mathbb{R}_+, L^2(R^{n-1}\times]0,T[)).$$

We also have

$$\begin{aligned}\sum_{j,L}\frac{\partial}{\partial x_n}a_{nj;KL}\frac{\partial u_L}{\partial x_j} &= -\sum_{i=1}^{n-1}\frac{\partial}{\partial x_i}a_{ij;KL}\frac{\partial u_L}{\partial x_j} + \sum_L c_{KL}\frac{\partial^2 u_L}{\partial t^2} + f_K\\ &\in L^2(\mathbb{R}_+, \bar{H}_{(1)}(\mathbb{R}^{n-1}\times]0,T[)).\end{aligned}$$

In other words, $\sum_{j,L} a_{nj;KL}\frac{\partial u_L}{\partial x_j}$ is an element of the anisotropic Sobolev space $H_{(1)}(\mathbb{R}^+, \bar{H}_{(-1)}(\mathbb{R}^{n-1}\times]0,T[))$, one can consult for instance Hörmander [26], appendix B for more information on such spaces. In particular $\sum_{j,L} a_{nj;KL}\frac{\partial u_L}{\partial x_j}$ is a continuous function of x_n with values in $\bar{H}_{(-1)}(\mathbb{R}^{n-1}\times]0,T[)$.

When the PDE (2.1) is satisfied (with sufficiently smooth u_L and $a_{ij;KL}$) and the boundary conditions are satisfied it follows using partial integration that the partial differential equation in weak form (2.8) is satisfied. By testing with $v \in C_0^\infty([0,T], \dot{H}_{(1)})$ and partial integration, (2.8) implies that the PDE is satisfied.

It remains to be shown that (2.8) also implies the boundary condition. To do this we use a sequence of test functions $v_m = \chi_m(x_n)\phi_K(x',t)$, where $\phi_K \in C_0^\infty(\mathbb{R}^{n-1} \times [0,T])$. We define $\chi_1(x_n) = 1$ for $x_n \leq 0$, $\chi_1(x_n) = 0$ for $x_n \geq 1$ and smoothly decreasing between $x_n = 0$ and $x_n = 1$. Let $\chi_m(x_n) = \chi_1(mx_n)$. Equation (2.8) tested with v_m gives

$$\begin{aligned}\int_{x_n>0}\Bigg(&\sum_{j,K,L} a_{nj;KL}\frac{\partial\chi_m}{\partial x_n}\overline{\phi_K}\frac{\partial u_L}{\partial x_j} + \sum_{i,j,K,L} a_{ij;KL}\chi_m\frac{\partial\overline{\phi_K}}{\partial x_i}\frac{\partial u_L}{\partial x_j} - \sum_{K,L} c_{KL}\chi_m\overline{\phi_K}u_L\\ &-\sum_K \chi_m\overline{\phi_K}f_K\Bigg)\,\mathrm{d}x\,\mathrm{d}t = 0. \qquad (2.14)\end{aligned}$$

When we take the limit $m \to \infty$ the second, third and fourth term converge to 0. The first term converges to

$$\int_{x_n=0}\sum_{j,K,L}\phi_K(x',t)a_{nj;KL}\frac{\partial u_L}{\partial x_j}\,\mathrm{d}x'\,\mathrm{d}t.$$

Since ϕ_K is arbitrary it follows that $\sum_{j,L} a_{nj;KL} \frac{\partial u_L}{\partial x_j}|_{x_n=0} = 0$. Note that by testing with $\phi_K(x',t)\chi_m(x_n - s)$ one can show that $\sum_{j,L} a_{nj;KL} \frac{\partial u_L}{\partial x_j}$ is a continuous function of $x_n = s$ with values in a space of distributions.

In Fattorini [20] the more general boundary condition with γu on the right hand side of (2.12) is discussed, where γ is some L^∞ function on the boundary ∂X. In this case the sesquilinear form $a(u,u) + \lambda\|u\|^2$ has to be generalized to include an integral over the boundary, see equation (4.2) in [20]. For this case the estimate (2.10) is obtained in equation (4.7) in [20].

2.4 Well-posedness for general second order hyperbolic equations

Motivated by the examples in the previous sections we will study the existence, uniqueness and stability of solutions to (2.8). The setup is that of Lions and Magenes [36], Section III.8. However, our presentation is somewhat different. We start by giving properties of solutions of (2.8), in particular that the solutions satisfy energy estimates, and that they are continuous with respect to time. Using these properties it follows easily that the solutions are unique and depend continuously on initial values and the source f. Existence of solutions is shown separately.

Let V and H be two Hilbert spaces, V dense in H, let $\|\ \ \|$ denote the norm in V, let $|\ \ |$ denote the norm in H and $\langle\ ,\ \rangle$ the sesquilinear scalar product in H. We identify H with its antidual, then $V \subset H \subset V'$. If $f \in V'$, $v \in V$ then $\langle f, v\rangle$ will denote their scalar product in the antiduality.

Let $A : V \to V'$ and $C : H \to H$ be continuous and symmetric operators. The operators A and C define sesquilinear forms $a(u,v) = \langle Au, v\rangle = \langle u, Av\rangle$ and $c(u,v) = \langle Cu, v\rangle = \langle u, Cv\rangle$ on V and H respectively. We impose a few conditions on a, c. We assume that a is positive definite in the sense that for some λ

$$a(u,u) + \lambda|u|^2 \geq \alpha\|u\|^2, \qquad \alpha > 0. \tag{2.15}$$

For c we assume

$$c(u,u) \geq \beta|u|^2, \qquad \beta > 0. \tag{2.16}$$

It follows that $A + \lambda$ and C are invertible operators.

Let u be an element of $L^2(]0,T[,V)$ such that $u' \in L^2(]0,T[,H)$. We will consider the following differential equation

$$Au(t) + Cu''(t) = f(t). \tag{2.17}$$

Equation (2.17) will be understood in the following weak sense

$$\int_0^T [a(u(t),v(t)) - c(u'(t),v'(t))]\,\mathrm{d}t = \int_0^T \langle f(t), v(t)\rangle\,\mathrm{d}t. \tag{2.18}$$

for all v in $C_0^\infty([0,T],V)$.

The initial conditions are

$$\begin{aligned} u(0) &= u_0, \\ u'(0) &= u_1, \end{aligned} \tag{2.19}$$

with $u_0 \in V$, $u_1 \in H$. Since $t \mapsto u(t)$ is only L^2, in general $u(0)$ is not defined in V. But we are considering u such that $u' \in L^2(]0,T[,H)$. From this it follows that (see [25], Section 4.5)

$$u(t) \text{ is Hölder continuous of order } \tfrac{1}{2} \text{ with values in } H, \tag{2.20}$$

so $u(0)$ is well defined in H. It follows from the differential equation that $u'' \in L^2(]0,T[,V')$. Hence

$$u'(t) \text{ is Hölder continuous of order } \tfrac{1}{2} \text{ with values in } V', \tag{2.21}$$

and $u'(0)$ is well defined in V'. Therefore equations (2.19) make sense.

We will first show that solutions to (2.18) satisfy energy estimates, where the energy is given by

$$E(t) = \tfrac{1}{2}\big(a(u(t),u(t)) + c(u'(t),u'(t)) + \lambda |u(t)|^2\big).$$

Lemma 2.4.1 *Suppose a, c are symmetric sesquilinear forms on V, H respectively that obey (2.15), (2.16). Let $u \in L^2(]0,T[,V), u' \in L^2(]0,T[,H), f \in L^2(]0,T[,H)$. If u satisfies (2.18), then there is a subset I_0 of $]0,T[$ that differs from $]0,T[$ by a set of measure zero, such that the energy $E(t) = \frac{1}{2}\left(a(u(t),u(t)) + c(u'(t),u'(t)) + \lambda|u|^2\right)$ is Hölder continuous of order $\frac{1}{2}$ on I_0 and it obeys*

$$E(t) \le C\left(E(s) + \int_0^T |f(\sigma)|^2 \,\mathrm{d}\sigma\right), \tag{2.22}$$

for all $s, t \in I_0$, $C = C(\alpha, \beta, \lambda, T)$.

Proof Let $\rho_m, m = 1, 2, \ldots$ be a regularizing sequence, that is

$$\begin{aligned} &\rho_m(t) \ge 0 \\ &\int \rho_m(t)\,\mathrm{d}t = 1 \\ &\operatorname{supp} \rho_m \to \{0\}. \end{aligned} \tag{2.23}$$

Let $\epsilon > 0$. There is an M such that $\operatorname{supp} \rho_m \subset]-\epsilon, \epsilon[$ for all $m > M$.

Set $v(t) = \rho_m(s-t)(\rho_m' \star u)(s)$, then, for $s \in [\epsilon, T-\epsilon]$, we have that $v \in C_0^\infty([0,T],V)$. Hence u satisfies (2.18) with this choice of v. But the integral $\int_0^T \rho_m(s-t)u(t)\,\mathrm{d}t$ simply equals the convolution product $\rho_m \star u(s)$, hence

$$a(\rho_m \star u(s), \rho_m' \star u(s)) + c(\rho_m'' \star u(s), \rho_m' \star u(s)) = \langle \rho_m \star f(s), \rho_m' \star u(s) \rangle.$$

We will use the following shorthand notation

$$\begin{aligned}
u_m(s) &= \rho_m \star u(s), \qquad f_m(s) = \rho_m \star f(s), \\
E_m(s) &= \tfrac{1}{2}\left(a(u_m(s), u_m(s)) + c(u'_m(s), u'_m(s)) + \lambda|u|^2\right),
\end{aligned}$$

Then it follows that

$$\frac{dE_m}{ds}(s) = \mathrm{Re}(\langle f_m(s), u'_m(s)\rangle + \lambda\langle u_m(s), u'_m(s)\rangle). \tag{2.24}$$

Define $u(t) = 0$ for t outside $[0, T]$, then we can define $u_m(t), u'_m(t), f_m(t)$ for all $t \in]0, T[$. In the limit $m \to \infty$ we have

$$\begin{aligned}
u_m &\to u && \text{in } L^2(]0, T[, V), \\
u'_m &\to u' && \text{in } L^2(]0, T[, H), \\
f_m &\to f && \text{in } L^2(]0, T[, H).
\end{aligned}$$

Hence

$$\begin{aligned}
E_m &\to E && \text{in } L^1(]0, T[), \\
\langle f_m, u'_m\rangle &\to \langle f, u'\rangle && \text{in } L^1(]0, T[), \\
\langle u_m(s), u'_m(s)\rangle &\to \langle u(s), u'(s)\rangle && \text{in } L^1(]0, T[).
\end{aligned}$$

Since ϵ is arbitrary small, equation (2.24) becomes

$$\frac{dE}{dt} = \mathrm{Re}(\langle f, u'\rangle + \lambda\langle u(s), u'(s)\rangle), \tag{2.25}$$

in $L^1(]0, T[)$ and we find that $\frac{dE}{dt}$ is in $L^1(]0, T[)$. It follows that the map $t \mapsto E(t)$ is absolutely continuous on $]0, T[$.

Hence there is a subset I_0 of $]0, T[$ that differs from $]0, T[$ by a set of measure zero, and a subsequence $m(l)$ such that for all $t \in I_0$ we have

$$\begin{aligned}
&u_{m(l)}(t) \to u(t) \text{ in } V, \qquad u'_{m(l)}(t) \to u'(t) \text{ in } H, \qquad f_{m(l)}(t) \to f(t) \text{ in } H, \\
&E_{m(l)}(t) \to E(t) \text{ in } \mathbb{R}, \qquad \frac{dE_{m(l)}}{dt}(t) \to \frac{dE}{dt}(t) \text{ in } \mathbb{R},
\end{aligned} \tag{2.26}$$

and (2.25).

We show that $E(t)$ is Hölder continuous of order $\frac{1}{2}$. Let $t \in I_0$. Since $\frac{dE}{dt} \in L^1(]0, T[)$, the energy $E(t)$ is bounded. This implies (see (2.15), (2.16)) that $\|u(t)\|$ and $|u'(t)|$ are bounded. It follows from (2.25) that (we will denote the various constants by C, C_1, C_2)

$$\frac{dE}{dt} \leq |f(t)|\,|u(t)| + \lambda|u(t)|\,|u'(t)| \leq C_1|f(t)| + C_2.$$

Integrating this inequality we obtain, for $s < t \in I_0$,

$$\begin{aligned}|E(t) - E(s)| &\le \int_s^t (C_1|f(\sigma)| + C_2)\, \mathrm{d}\sigma \\ &= C_1 \int_0^T 1_{[s,t]}\, |f(\sigma)|\, \mathrm{d}\sigma + C_2(t-s) \\ &\le C_1 \|1_{[s,t]}\|_{L^2(]0,T[)} \cdot \|f\|_{L^2(]0,T[,H)} + C_2(t-s) \\ &= C_1\sqrt{t-s}\, \left[\int_0^T |f(\sigma)|^2\, \mathrm{d}\sigma\right]^{1/2} + C_2(t-s). \end{aligned} \tag{2.27}$$

This shows that $E(t)$ is Hölder continuous of order $\frac{1}{2}$ on I_0.

Finally we prove (2.22). Let $s, t \in I_0$. We obtain by integrating (2.24)

$$E_m(t) \le E_m(s) + \int_s^t \tfrac{1}{2}(|f_m(\sigma)|^2 + (1+\lambda)|u_m'(\sigma)|^2 + \lambda|u_m|^2)\, \mathrm{d}\sigma.$$

Using (2.15), (2.16) we obtain

$$E_m(t) \le E_m(s) + \int_0^T |\tfrac{1}{2} f(\sigma)|^2\, \mathrm{d}\sigma + C\int_s^t E_m(\sigma)\, \mathrm{d}\sigma,$$

C a constant depending on α, β, λ. Hence, using Gronwall's lemma we find

$$E_m(t) \le C\left(E_m(s) + \int_0^T \tfrac{1}{2}|f(\sigma)|^2\, \mathrm{d}\sigma\right), \tag{2.28}$$

C depending on $\alpha, \beta, \lambda, T$. Taking the limit for $m \to \infty$ we obtain (2.22). $\square$

Since the energy is bounded almost everywhere on $]0,T[$ it follows from (2.15) and (2.16) that a solution u of (2.18) satisfies

$$u \in L^\infty(]0,T[, V), \qquad u' \in L^\infty(]0,T[, H). \tag{2.29}$$

Using this we are now going to prove that u and u' are uniformly continuous in V and H respectively. This implies that they can be defined on the closed interval $[0,T]$. Here we follow Strauss [54].

Theorem 2.4.2 *Let the assumptions of Lemma 2.4.1 be satisfied. Then, after modification of the function $t \mapsto u(t)$ on a set of measure zero, we have*

$$u \in C([0,T], V), \qquad u' \in C([0,T], H). \tag{2.30}$$

The proof requires a few lemmas. We define a function u from a metric space I to a Banach space E to be weakly continuous (resp. weakly uniformly continuous) if the map $t \mapsto \langle u(t), v\rangle$ is continuous (resp. uniformly continuous) on I for all $v \in E'$.

Lemma 2.4.3 *Let I be a metric space, provided with a Borel measure, for which every nonvoid open subset has positive measure. Let E be a Banach space, $u \in L^\infty(I, E)$, D a dense subset of E' with respect to the operator norm in E', and for every $v \in D$ the map $t \mapsto \langle u(t), v\rangle$ is uniformly continuous. Then, after modification on a set of measure zero, the function $t \mapsto u(t)$ is weakly uniformly continuous, which means that for every $v \in E'$ the mapping $t \mapsto \langle u(t), v\rangle$ is uniformly continuous.*

Proof Let $I_0 = \{t \in I | \, \|u(t)\| \leq \|u\|_{L^\infty}\}$, then $I \backslash I_0$ has measure zero, hence I_0 is dense in I, because of the assumptions on I. Let $v \in E'$ and $\epsilon > 0$. There exists $w = w(v, \epsilon) \in D$ such that

$$\|v - w\| \cdot \|u\|_{L^\infty} < \frac{\epsilon}{2}.$$

Furthermore, there exists $\delta > 0$ such that

$$t, s \in I, d(t, s) < \delta \quad \Rightarrow \quad |\langle u(t), w\rangle - \langle u(s), w\rangle| < \frac{\epsilon}{2}.$$

Hence, for $t, s \in I_0$

$$|\langle u(t), v\rangle - \langle u(s), v\rangle| \leq |\langle u(t), w\rangle - \langle u(s), w\rangle| + |\langle u(t), v - w\rangle - \langle u(s), v - w\rangle| < \epsilon.$$

It follows that $t \mapsto \langle u(t), v\rangle$ is uniformly continuous on I_0 hence on the completion of I_0, which contains I. □

Lemma 2.4.4 *Under the conditions of Lemma 2.4.1, after modification on a set of measure zero, the solution u of (2.18) (resp. its derivative u') is a weakly uniformly continuous mapping $[0, T] \to V$ (resp. $[0, T] \to H$).*

Proof We will use Lemma 2.4.3 with $I = [0, T]$.

Take $E = V, D = H$. We know from (2.29) that $u \in L^\infty([0, T], V)$. By (2.20) u is Hölder continuous of order $\frac{1}{2}$ with values in H. In particular the map $t \mapsto \langle u(t), v\rangle$ is uniformly continuous for $v \in H$. Hence by the lemma the map $t \mapsto \langle u(t), v\rangle$ is uniformly continuous for $v \in V'$, after modification on a set of measure zero.

Now take $E = H, D = V$. Again from (2.29) we know that $u' \in L^\infty([0, T], H)$ and by (2.21) u' is Hölder continuous of order $\frac{1}{2}$ with values in V'. In particular the map $t \mapsto \langle u'(t), v\rangle$ is uniformly continuous for $v \in V$. Hence by the lemma the map $t \mapsto \langle u(t), v\rangle$ is uniformly continuous for $v \in H$, after modification on a set of measure zero. □

Proof of Theorem 2.4.2 There is a subset I_0 of $[0,T]$, that differs from $[0,T]$ by a set of measure zero, such that the map $t \mapsto E(t)$ is uniformly continuous on I_0, $u(t) \in V$, $u'(t) \in H$ for $t \in I_0$ and $t \mapsto u(t)$, $t \mapsto u'(t)$ are weakly uniformly continuous maps on I_0.

Let $s, t \in I_0$ and

$$\xi = \tfrac{1}{2}\left(a(u(s)-u(t), u(s)-u(t)) + c(u'(s)-u'(t), u'(s)-u'(t)) + \lambda|u(s)-u(t)|^2\right).$$

We have

$$\xi = E(s) + E(t) - a(u(s), u(t)) - c(u'(s), u'(t)) - \lambda\langle u(s), u(t)\rangle.$$

Now take the limit $s \to t$. By Lemma 2.4.1 the first term converges uniformly to $E(t)$. By Lemma 2.4.4 the third and fourth terms converge uniformly to $-a(u(t), u(t))$ and $-c(u'(t), u'(t))$ respectively. Because of (2.20) the last term converges uniformly to $-\lambda|u(t)|^2$. Hence

$$\xi \to E(t) + E(t) - a(u(t), u(t)) - c(u'(t), u'(t)) - \lambda|u(t)|^2 = 0,$$

uniformly for $s, t \in I_0$. Since $2\xi \geq \alpha\|u(s)-u(t)\|^2 + \beta|u'(s)-u'(t)|^2$, it follows that the mapping $t \mapsto u(t)$ is uniformly continuous from I_0 to V hence on its completion $[0,T]$. Similarly the mapping $t \mapsto u'(t)$ is uniformly continuous from I_0 to H hence on its completion $[0,T]$. This proves the theorem. □

We can now prove well-posedness of the problem, i.e. existence, uniqueness and continuous dependence of the solution on the parameters.

Theorem 2.4.5 *Suppose that a is a sesquilinear form on V that obeys (2.15), c is a sesquilinear form on H that obeys (2.16). Suppose*

$$u_0 \in V, \qquad u_1 \in H, \qquad f \in L^2(]0,T[, H).$$

Then the problem (2.18), (2.19) has a solution

$$u \in C([0,T], V) \cap C^1([0,T], H). \tag{2.31}$$

The solution u is unique in $L^2(]0,T[, V) \cap \bar{H}_{(1)}(]0,T[, H)$ and depends continuously on u_0, u_1 and f.

Proof of the existence The proof of the existence proceeds similarly as in Lions and Magenes [36].

To simplify matters we will assume that V is separable, which will be the case in all our applications. Let $w_1, \dots, w_m, \dots$ form a "basis" for V in the following sense:

for all m, $w_1, \dots, w_m$ are linearly independent, and,

the combinations $\sum_{\text{finite}} \xi_k w_k, \xi_k \in \mathbb{R}$, are dense in V.

Such a basis always exists if V is separable.

We define the approximate solution $u_m(t)$ of order m of the problem in the following way

$$u_m(t) = \sum_{k=1}^{m} g_{km}(t) w_k,$$

the g_{km}'s being determined so that

$$a(u_m(t), w_l) + c(u_m''(t), w_l) = (f(t), w_l), \qquad 1 \le l \le m, \tag{2.32}$$

$$g_{km}(0) = \xi_{km}, \qquad g_{km}'(0) = \eta_{km} \tag{2.33}$$

with

$$\sum_{k=1}^{m} \xi_{km} w_k \to u_0 \qquad \text{in } V \text{ as } m \to \infty,$$
$$\sum_{k=1}^{m} \eta_{km} w_k \to u_1 \qquad \text{in } H \text{ as } m \to \infty. \tag{2.34}$$

Thus, the g_{km}'s are determined by a linear differential system which admits a unique solution. The same calculations as in the section on energy estimates above show that

$$\|u_m(t)\|^2 + |u_m'(t)|^2 \le \text{ constant depending on } f, u_0, u_1, a, c.$$

Therefore, in particular, u_m (resp. u_m') remains in a bounded set of $L^2(]0,T[,V)$ (resp. $L^2(]0,T[,H)$). By (2.32), it follows that u_m'' stays in a bounded set of $L^2(]0,T[,V')$. Since the unit ball is weakly compact in a reflexive Banach space, we may extract $m(l)$, so that [18, II.3.26]

$$u_{m(l)} \to u \quad \text{weakly in } L^2(]0,T[,V),$$
$$u_{m(l)}' \to \chi \quad \text{weakly in } L^2(]0,T[,H) \tag{2.35}$$

as $l \to \infty$. But $\chi = u'$.

It remains to be shown that the function u constructed in this manner is a solution of the problem (2.18), (2.19). We first show that u satisfies the differential equation. To this end we consider the functions

$$\psi = \sum_{k=1}^{m_0} \phi_k \otimes w_k, \qquad \phi_k \in C_0^\infty([0,T]). \tag{2.36}$$

For $m > m_0$, we deduce from (2.32) (multiplying by $\phi_k(t)$ and summing over k from 1 to m_0) that

$$\int_0^T [a(u_{m(l)}, \psi) - c(u_{m(l)}', \psi')] \,\mathrm{d}t = \int_0^T \langle f, \psi \rangle \,\mathrm{d}t.$$

By taking the limit $l \to \infty$ it follows from this that, for all ψ of the form given above,

$$\int_0^T [a(u, \psi) - c(u', \psi')] \, \mathrm{d}t = \int_0^T \langle f, \psi \rangle \, \mathrm{d}t.$$

But since the w_m's form a basis for V, the set of functions of the form (2.36) is dense in $L^2(]0, T[, V)$. Hence u satisfies the differential equation.

We show that u satisfies the first inital condition. Now define the following space of functions

$$C_T^\infty = \{\phi | \phi \in C^\infty([0, T]), \phi(T) = \phi'(T) = 0\}, \tag{2.37}$$

and consider the functions of the form

$$\psi = \phi \otimes v, \qquad v \in V, \phi \in C_T^\infty([0, T]). \tag{2.38}$$

By partial integration we have

$$\int_0^T [(u_m', \psi) + (u_m, \psi')] \, \mathrm{d}t = -(u_m(0), \psi(0)).$$

The right hand side converges to $(u_0, \psi(0))$ if $m \to \infty$. The left hand side converges to

$$\int_0^T [(u', \psi) + (u, \psi')] \, \mathrm{d}t = (u(0), \psi(0)).$$

Hence

$$(u(0), \psi(0)) = (u_0, \psi(0)). \tag{2.39}$$

Using (2.33) it follows that u satisfies the first initial condition.

The differential equation, tested with ψ as in (2.38) has an extra term due to the partial integration

$$\int_0^T [a(u_{m(l)}, \psi) - c(u_{m(l)}', \psi')] \, \mathrm{d}t = \int_0^T \langle f, \psi \rangle \, \mathrm{d}t + c(u_{m(l)}'(0), \psi(0)).$$

Let $\psi_k(t) = \chi_k(t) \otimes v$, $v \in V$, χ_k as defined above (2.14). By choosing k sufficiently large the first term can be made arbitrarily small, and (because u is continuous by Theorem 2.4.2) $\int_0^T c(u'(t), \psi_k'(t)) \, \mathrm{d}t - c(u'(0), \psi(0))$ can be chosen arbitrarily small. Letting $l \to \infty$ we find that $c(u_{m(l)}'(0), \psi(0)) \to c(u'(0), \psi(0))$. Using (2.33) it follows that u satisfies the second initial condition. $\square$

Proof of uniqueness and continuous dependence on parameters Suppose $\tilde{u}$ and $\bar{u}$ both satisfy (2.18), (2.19), then their difference $u = \tilde{u} - \bar{u}$ satisfies the differential equation with $u_0 = u_1 = 0$, $f = 0$.

By Theorem 2.4.2 $u(t)$ is continuous in V and $u'(t)$ is continuous in H. Hence we can use estimate (2.22) with $s = 0$, t arbitrary in $[0, T]$,

$$\begin{aligned} E(t) &\leq C\left(E(0) + \int_0^T |f(\sigma)|^2 \,\mathrm{d}\sigma\right) \\ &= C\left(\|u_0\|^2 + |u_1|^2 + \int_0^T |f(\sigma)|^2 \,\mathrm{d}\sigma\right) = 0. \end{aligned}$$

Hence by (2.15), (2.16) we have $u = 0$, and $\tilde{u} = \bar{u}$.

Suppose now that we have two neighboring problems $\tilde{u}_0, \tilde{u}_1 \tilde{f}$ and $\bar{u}_0, \bar{u}_1, \bar{f}$. Then $u = \tilde{u} - \bar{u}$ satisfies the differential equation with $u_0 = \tilde{u}_0 - \bar{u}_0$, $u_1 = \tilde{u}_1 - \bar{u}_1$ and $f = \tilde{f} - \bar{f}$. By (2.22), with $s = 0$, t arbitrary in $[0, T]$ we have

$$E(t) \leq C\left(\|\tilde{u}_0 - \bar{u}_0\|^2 + |\tilde{u}_1 - \bar{u}_1|^2 + \int_0^T |\tilde{f}(\sigma) - \bar{f}(\sigma)|^2 \,\mathrm{d}\sigma\right).$$

Therefore $\tilde{u} \to \bar{u}$ if $\tilde{u}_0 \to \bar{u}_0$, $\tilde{u}_1 \to \bar{u}_1$ and $\tilde{f} \to \bar{f}$ in V, H and $L^2(]0, T[, H)$ respectively. $\square$

2.5 More general differential equations

In this section we will consider a generalization of (2.1), with time dependent coefficients and lower order terms added

$$\begin{aligned} \sum_L \Bigg(-\sum_{i,j} \tfrac{\partial}{\partial x_i} a_{ij;KL}(x,t) \tfrac{\partial}{\partial x_j} + \sum_i a^{(1)}_{i;KL}(x,t) \tfrac{\partial}{\partial x_i} + a^{(2)}_{KL}(x,t) + b_{KL}(x,t) \tfrac{\partial}{\partial t} \\ + \tfrac{\partial}{\partial t} c_{KL}(x,t) \tfrac{\partial}{\partial t} \Bigg) u_L(x,t) = f_K(x,t). \end{aligned} \tag{2.40}$$

This equation can be written in abstract form similarly to (2.8), and the techniques of the previous section can be applied to this equation. The solutions satisfy energy estimates, and the other results (continuity of the solutions, existence, uniqueness) follow.

Equation (2.40) can be written in the following form

$$\begin{aligned} \int_0^T [a(t; u(t), v(t)) + a_1(t; u(t), v(t)) + b(t; u'(t), v(t)) - c(t; u'(t), v'(t))] \,\mathrm{d}t \\ = \int_0^T \langle f(t), v(t) \rangle \,\mathrm{d}t. \end{aligned} \tag{2.41}$$

for all $v \in C_0^\infty([0,T],V)$. The symmetric sesquilinear form a contains terms involving the coefficient $a(x,t)$ and the symmetric part of $a^{(2)}(x,t)$. The sesquilinear form a_1 (in general not symmetric) contains the remaining terms.

We assume $a(t), c(t)$ are symmetric sesquilinear forms satisfying (2.15) and (2.16) respectively for each t with constants α, β, λ independent of t. We assume they only have weak time dependence of the form

the map $t \mapsto a(t)$ from $[0,T]$ to the space of continuous sesquilinear forms on V is once continuously differentiable, (2.42)

and

the map $t \mapsto c(t)$ from $[0,T]$ to the space of continuous sesquilinear forms on H is once continuously differentiable. (2.43)

The derivatives are denoted $a'(t;u,v)$ and $c'(t;u,v)$. This excludes for instance moving discontinuities in the coefficients. The sesquilinear form $a_1(t;u,v)$ is defined for $u \in V, v \in H$, while b is a sesquilinear form on H. We require for the time dependence of a_1, b

the map $t \mapsto a_1(t)$ from $[0,T]$ to the space of continuous sesquilinear forms on $V \times H$ is continuous,

the map $t \mapsto b(t)$ from $[0,T]$ to the space of continuous sesquilinear forms on H is continuous, (2.44)

We will show the following energy estimate.

Lemma 2.5.1 *Suppose $a(t;u,v)$ is a family of sesquilinear forms on V that obeys (2.42) and (2.15) for each t. Suppose $c(t;u,v)$ is a family of sesquilinear forms on H that obeys (2.43) and (2.16) for each t. Suppose $a_1(t;u,v), b(t;u,v)$ are families of bilinear forms with $u \in V, v \in H$ or $u,v \in H$ respectively, that obey (2.44). Let $f \in L^2(]0,T[,H)$. Suppose $u \in L^2(]0,T[,V)$, with $u' \in L^2(]0,T[,H)$, satisfies the differential equation (2.41). Then there is a subset I_0 of $]0,T[$ that differs from $]0,T[$ by a set of measure zero, such that the energy*

$$E = \tfrac{1}{2}[a(t;u(t),u(t)) + \lambda|u(t)|^2 + c(t;u'(t),u'(t))]$$

is Hölder continuous of order $\frac{1}{2}$ on I_0 and it obeys

$$E(t) \leq C[E(s) + \int_0^T |f(\sigma)|^2 \, \mathrm{d}\sigma], \tag{2.45}$$

for all $s,t \in I_0$, $C = C(a,a_1,b,c,T)$.

Proof Let $\rho_m, m = 1, 2, \ldots$ be a regularizing sequence (see (2.23)). Let $\epsilon > 0$. There is an M such that $\operatorname{supp} \rho_m \subset]-\epsilon, \epsilon[$ for all $m > M$.

Set $v(t) = \rho_m(s-t)(\rho'_m \star u(s))$, then, for $s \in [\epsilon, T-\epsilon]$, we have that $v \in C_0^\infty([0,T], V)$. Hence u satisfies (2.41) with this choice of v:

$$\begin{aligned}\int_0^T &[a(t; u(t), \rho_m(s-t)(\rho'_m \star u(s)) + a_1(t; u(t), \rho_m(s-t)(\rho'_m \star u(s)) \\ &\quad + b(t; u'(t), \rho_m(s-t)(\rho'_m \star u(s)) + c(t; u'(t), \rho'_m(s-t)(\rho'_m \star u(s))] \, \mathrm{d}t \\ &= \int_0^T \langle f(t), \rho_m(s-t)(\rho'_m \star u(s))\rangle \, \mathrm{d}t. \end{aligned} \tag{2.46}$$

The first term of (2.46) equals

$$\langle (\rho_m \star Au)(s), \rho'_m \star u(s)\rangle.$$

Since the operator A is time dependent is does not commute with the operator $\rho_m\star$. However we will show in a separate lemma that the commutator term (that I will call $R_{a,m}$) vanishes in the limit $m \to \infty$. So we write the previous expression as

$$a(s; \rho_m \star u(s), \rho'_m \star u(s)) + R_{a,m}(s),$$

where we defined $R_{a,m}(s)$ by

$$R_{a,m}(s) = \langle [\rho_m \star Au - A(\rho_m \star u)](s), \rho'_m \star u(s)\rangle.$$

Let $R_{a_1,m}$ be similarly defined for the second term in (2.46). The third term in (2.46) equals

$$b(s; \rho'_m \star u(s), \rho'_m \star u(s)) + R_{b,m}(s),$$

with

$$R_{b,m}(s) = \langle [\rho_m \star Bu' - B(\rho_m \star u')](s), \rho'_m \star u(s)\rangle.$$

The fourth term in (2.46) equals

$$c(s; \rho''_m \star u(s), \rho'_m \star u(s)) + R_{c,m}(s),$$

with

$$R_{c,m}(s) = \langle [\rho'_m \star Cu' - C(\rho'_m \star u')](s), \rho'_m \star u(s)\rangle.$$

In Lemma 2.5.2 it is shown that $R_{a,m}, R_{a_1,m}, R_{b,m}, R_{c,m}$ vanish in the limit $m \to \infty$.

We will use the following shorthand notation

$$\begin{aligned} u_m(s) &= \rho_m \star u(s), \qquad\qquad f_m(s) = \rho_m \star f(s), \\ E_m(s) &= \tfrac{1}{2}[a(s; u_m(s), u_m(s)) + \lambda |u_m(s)|^2 + c(s; u'_m(s), u'_m(s))]. \end{aligned}$$

Then this can be written

$$
\begin{aligned}
&a(s; u_m(s), u'_m(s)) + a_1(s; u_m(s), u'_m(s)) + b(s; u'_m(s), u'_m(s)) + c(s; u''_m(s), u'_m(s)) \\
&\quad = \langle f_m(s), u'_m(s)\rangle - R_{a,m}(s) - R_{a_1,m} - R_{b,m}(s) - R_{c,m}(s).
\end{aligned}
$$

Since

$$
\begin{aligned}
2\,\mathrm{Re}\, a(s; u_m(s), u'_m(s)) &= a(s; u_m(s), u'_m(s)) + a(s; u'_m(s), u_m(s)) \\
&= \frac{d}{ds} a(s; u_m(s), u_m(s)) - a'(s; u_m(s), u_m(s))
\end{aligned}
$$

(and similarly for $c(t; u_m(t), u_m(t))$), we have

$$
\begin{aligned}
\frac{dE_m}{ds} = \; &\mathrm{Re}(\langle f_m(s), u'_m(s)\rangle - a_1(s; u_m(s), u'_m(s)) - b(s; u'_m(s), u'_m(s)) \\
&+ \tfrac{1}{2} a'(s; u_m(s), u_m(s)) + \tfrac{1}{2} c'(s; u'_m(s), u'_m(s)) + \lambda \langle u_m(s), u'_m(s)\rangle \\
&- R_{a,m}(s) - R_{a_1,m}(s) - R_{b,m}(s) - R_{c,m}(s)).
\end{aligned}
\tag{2.47}
$$

Define $u(t) = 0$ for t outside $[0, T]$, then we can define $u_m(t), u'_m(t), f_m(t)$ and $R_{a,m}(t)$, $R_{a_1,m}(t)$, $R_{b,m}(t)$, $R_{c,m}(t)$ for all $t \in]0, T[$. In the limit $m \to \infty$ we have

$$
\begin{aligned}
u_m &\to u && \text{in } L^2(]0, T[, V), \\
u'_m &\to u && \text{in } L^2(]0, T[, H), \\
f_m &\to f && \text{in } L^2(]0, T[, H).
\end{aligned}
$$

And, by Lemma 2.5.2,

$$
R_{a,m}, R_{a_1,m}, R_{b,m}, R_{c,m} \to 0 \text{ in } L^1(]0, T[).
$$

Since ϵ is arbitrary we can take the limit $m \to \infty$ in (2.47). We obtain

$$
\begin{aligned}
\frac{dE}{dt}(t) = \; &\mathrm{Re}(\langle f(t), u'(t)\rangle - a_1(t; u(t), u'(t)) - b(t; u'(t), u'(t)) + \tfrac{1}{2} a'(t; u(t), u(t)) \\
&+ \tfrac{1}{2} c'(t; u'(t), u'(t)) + \lambda \langle u(t), u'(t)\rangle),
\end{aligned}
\tag{2.48}
$$

in $L^1(]0, T[)$. So the map $t \mapsto E(t)$ is absolutely continuous on $]0, T[$.

It follows that there is a subset I_0 of $]0, T[$ that differs from $]0, T[$ by a set of measure zero, and a subsequence $m(l)$, such that for all $t \in I_0$ we have

$$
\begin{aligned}
&u_{m(l)}(t) \to u(t) \text{ in } V, && u'_{m(l)}(t) \to u'(t) \text{ in } H, && f_{m(l)}(t) \to f(t) \text{ in } H, \\
&E_{m(l)}(t) \to E(t) \text{ in } \mathbb{R}, && \frac{dE_{m(l)}}{dt}(t) \to \frac{dE}{dt}(t) \text{ in } \mathbb{R},
\end{aligned}
\tag{2.49}
$$

and such that (2.48) is valid pointwise.

We prove that the map $t \mapsto E(t)$ is Hölder continuous of order $\frac{1}{2}$. Let $t \in I_0$. The different terms of (2.48) can be estimated as follows (the various constants that arise are denoted by $C, C_1, C_2, \ldots$)

$$\begin{aligned}
&\langle f(t), u'(t)\rangle \leq |f(t)| \cdot |u'(t)|, && a'(t; u(t), u(t)) \leq C\,\|u(t)\|^2,\\
&a_1(t; u(t), u'(t)) \leq C\,\|u(t)\| \cdot |u'(t)| && b(t; u'(t), u'(t)) \leq C\,|u'(t)|^2,\\
&c'(t; u'(t), u'(t)) \leq C\,|u(t)|^2, && \langle u(t), u'(t)\rangle \leq\ |u(t)| \cdot |u'(t)|.
\end{aligned} \tag{2.50}$$

Since $E(t)$ is absolutely continuous it is bounded. This implies (see (2.15) and (2.16)) that $\|u(t)\|$, $|u'(t)|$ are bounded too, hence we have

$$\frac{dE}{dt}(t) \leq C_1 |f(t)| + C_2.$$

Integrating this equation we obtain in the same way as in (2.27), for $s, t \in I_0$

$$|E(t) - E(s)| \leq \sqrt{t-s}\, C_1 \left[\int_0^T |f(\sigma)|^2 \,\mathrm{d}\sigma\right]^{\frac{1}{2}} + (t-s)C_2.$$

This shows that $E(t)$ is Hölder continuous of order $\frac{1}{2}$ on I_0.

Finally we prove (2.45). Let $s, t \in I_0$. We obtain by integrating (2.47) and using estimates similar to (2.50)

$$E_m(t) \leq E_m(s) + \int_s^t \tfrac{1}{2}(|f_m(\sigma)|^2 + C_1\|u_m(\sigma)\|^2 + C_2|u_m'(\sigma)|^2)\,\mathrm{d}\sigma.$$

Using (2.15), (2.16) and Gronwall's lemma as in the derivation of (2.28) we find

$$E_m(t) \leq C\left(E_m(s) + \int_0^T \tfrac{1}{2}|f(\sigma)|^2 \,\mathrm{d}\sigma\right),$$

C depending on a, a_1, b, c, T. Taking the limit $m \to \infty$ we obtain (2.45). □

The following lemma completes the proof

Lemma 2.5.2 *Let a, a_1, b, c be as in Lemma 2.5.1 and $R_{a,m}, R_{a_1,m}, R_{b,m}, R_{c,m}$ as in the proof of Lemma 2.5.1. Then*

$$R_{a,m}, R_{a_1,m}, R_{b,m}, R_{c,m} \to 0 \qquad \text{in } L^1(]0,T[). \tag{2.51}$$

Proof We first show that $R_{a_1,m} \to 0$ in $L^1(]0,T[)$. The L^1 norm of $R_{a_1,m}$ can be estimated by using standard inequalities, such as

$$\begin{aligned}
\int |\, R_{a_1,m}(s)\,|\,\mathrm{d}s &= \int |\langle[\rho_m \star A_1 u - A_1(\rho_m \star u)](s), \rho_m' \star u(s)\rangle|\ \mathrm{d}s\\
&\leq \|\rho_m \star A_1 u - A_1(\rho_m \star u)\|_{L^2(]0,T[,H)} \cdot \|\rho_m' \star u\|_{L^2(]0,T[,H)}.
\end{aligned}$$

We prove that $\rho_m \star A_1 u - A_1(\rho_m \star u) \to 0$ in $L^2(]0,T[,H)$, then $R_{a_1,m} \to 0$ in $L^1(]0,T[)$. Let $s,t \in]0,T[$, then, by (2.44), $\|A_1(t) - A_1(s)\|_{V\to H} \to 0$ uniformly if $t \to s$. Hence for every $\epsilon > 0$ there is an M such that for $m > M$ we have

$$\rho_m(s-t)\,|((A_1(t) - A_1(s))u(t)| \leq \epsilon \rho_m(s-t)\|u(t)\|.$$

Hence, by a standard equality for the convolution product, $\rho_m \star A_1 u - A_1(\rho_m \star u) \to 0$ in $L^2(]0,T[,H)$.

By replacing $A_1 u$ by Bu' one can show similarly that also $R_{b,m} \to 0$ in $L^1(]0,T[)$.

We now show that $R_{a,m} \to 0$. Consider the map

$$T_m : v \mapsto \langle \rho_m \star Au - A(\rho_m \star)u, \rho_m' \star v \rangle,$$

from $L^2(]0,T[,V)$ to $L^1(]0,T[)$. We show that $T_m \to 0$ on a dense subspace of $L^2(]0,T[,V)$, and that the T_m are bounded, then (by Dunford & Schwartz, II.3.6) $T_m(v) \to 0$ for v in all of $L^2(]0,T[,V)$, which is what we have to show.

Choose a regularizing sequence of the form

$$\begin{aligned} &\rho_m(t) = \frac{1}{m}\rho(mt) \\ &\operatorname{supp}\rho = [-1,1], \qquad \rho > 0, \qquad \int \rho \,\mathrm{d}t = 1. \end{aligned}$$

It follows that

$$\int |\rho_m'(t)| \,\mathrm{d}t = C \cdot m.$$

Let $s,t \in]0,T[$, then, by (2.42), $\|A(t) - A(s)\|_{V\to V'} \leq C|s-t|$.

$$\begin{aligned} \|[\rho_m \star Au - A(\rho_m \star)u](s)\|_{V'} &\leq \int_{s-1/m}^{s+1/m} \rho_m(s-t)\|(A(t)-A(s))u(t)\|_{V'} \,\mathrm{d}t \\ &\leq Cm^{-1}(\rho_m \star \|u\|)(s). \end{aligned} \tag{2.52}$$

It follows that

$$\|\rho_m \star Au - A(\rho_m \star)u\|_{L^2(]0,T[,V')} \leq Cm^{-1}\|u\|_{L^2(]0,T[,V)}.$$

This implies that $T_m \to 0$ for v in $C^1(]0,T[,V)$, which is a dense subspace of $L^2(]0,T[,V)$.

We now show that T_m is bounded. The L^2 norm of v' can be estimated by

$$\begin{aligned} \|v_m'\|_{L^2(]0,T[,V)} &\leq \|v\|_{L^2(]0,T[,V)} \int |\rho_m'| \,\mathrm{d}t \\ &= Cm\|v\|_{L^2(]0,T[,V)}. \end{aligned} \tag{2.53}$$

Hence

$$\begin{aligned}\|T_m(v)\|_{L^1(]0,T[)} &= \int_0^T \langle[\rho_m \star Au - A(\rho_m\star)u](s), \rho_m' \star v(s)\rangle\,\mathrm{d}s \\ &\le \|\rho_m \star Au - A(\rho_m\star)u\|_{L^2(]0,T[,V')} \cdot \|v_m'\|_{L^2(]0,T[,V)} \\ &= C\|u\|_{L^2(]0,T[,V)} \cdot \|v\|_{L^2(]0,T[,V)}\end{aligned}$$

which shows that T_m is bounded.

Let us consider $R_{c,m}$. The idea of the proof will be the same as in the case of $R_{a,m}$. Define

$$t_m(v) = \rho_m' \star Cv' - C(\rho_m' \star v')$$

and consider the map

$$T_m : v \mapsto \langle t_m(v), \rho_m' \star u\rangle.$$

from $L^2(]0,T[,V)$ to $L^1(]0,T[)$. Then $R_{c,m} = T_m(u)$. We show that $T_m \to 0$ if $v \in C_0^\infty([0,T],V)$ (dense subspace of $L^2(]0,T[,V)$). We can write out the first factor in the inner product

$$t_m(v)(s) = \int \rho_m'(s-t)(C(t) - C(s))v'(t)\,\mathrm{d}t.$$

Since v is smooth we can partially integrate the derivative off ρ_m'. Hence

$$t_m(v)(s) = \int (\rho_m(s-t)(C(t) - C(s))v''(t) + \rho_m(s-t)C'(t)v'(t))\,\mathrm{d}t.$$

By an argument that is similar to the ones used above it follows that $t_m(v) \to 0$ in $L^2(]0,T[,H)$.

Remains to be shown that T_m is bounded. This can be done by estimates similar to (2.52), (2.53). □

Having proven the energy lemma for this case the other results of the previous section follow immediately. Indeed we have

$$u \in L^\infty(]0,T[,V), \qquad u' \in L^\infty(]0,T[,H),$$

and therefore Lemma 2.4.4 applies. Using this and the energy lemma we can prove Theorems 2.4.2, 2.4.5 for this case. The results are stated below. The proofs are omitted since it can be checked that the arguments used to prove Theorems 2.4.2, 2.4.5 are also valid in this case.

Theorem 2.5.3 *Suppose a, a_1, b, c are families of sesquilinear forms that satisfy the assumptions of Lemma 2.5.1. Let*

$$f \in L^2(]0,T[,H).$$

Let u satisfy (2.41), and $u \in L^2(]0,T[,V)$, $u' \in L^2(]0,T[,H)$. Then, after modification on a set of measure zero, we have

$$u \in C([0,T],V), \qquad\qquad u' \in C([0,T],H). \tag{2.54}$$

Theorem 2.5.4 *Suppose a, a_1, b, c are families of sesquilinear forms that satisfy the assumptions of Lemma 2.5.1. Suppose*

$$\begin{aligned} u_0 &\in V, & u_1 &\in H, \\ f &\in L^2(]0,T[,H). \end{aligned}$$

Then the problem (2.41), (2.19) has a solution

$$u \in C([0,T],V) \cap C^1([0,T],H). \tag{2.55}$$

The solution u is unique in the space of $u \in L^2(]0,T[,V)$ such that $u' \in L^2(]0,T[,H)$, and depends continuously on u_0, u_1 and f.

2.6 Solutions in larger or smaller spaces

Wave equations with smooth coefficients have well-defined solutions in Sobolev spaces $H_{(s)}$ for all orders $s \in \mathbb{R}$, see for example Hörmander [26], Chapter 23. We will discuss a similar result for nonsmooth coefficients. There is a sequence of spaces $V_{(k)}$, $k \in \mathbb{Z}$, such that the differential equation (2.17), with source $f \in L^2(]0,T[,V_{(k)})$ and initial values $u_0 \in V_{(k+1)}, u_1 \in V_{(k)}$ has a unique solution that is continuous in t with values in $V_{(k)}$.

To keep the argument simple we discuss the time independent problem without lower order terms of Section 2.4. We will assume that the operator C is the identity. This can be obtained by making the following transformation

$$\begin{aligned} u|_{\text{new}} &= C^{\frac{1}{2}} u|_{\text{old}}, & f|_{\text{new}} &= C^{-\frac{1}{2}} f|_{\text{old}}, \\ A|_{\text{new}} &= C^{-\frac{1}{2}} A C^{-\frac{1}{2}}|_{\text{old}}, & V|_{\text{new}} &= \{u \in H \mid C^{-\frac{1}{2}} u \in V\}. \end{aligned} \tag{2.56}$$

Because C is positive definite the operator $C^{\frac{1}{2}}$ can be constructed using the spectral decomposition of the operator C. Note that now the space V depends on C. (For this reason we did not discuss this simplified system so far.)

A space that is smaller than V is given by the domain $D(A)$ of A

$$D(A) = \{u \in V \mid Au \in H\}.$$

With norm given by the norm of the graph

$$\|u\|^2_{D(A)} = |Au|^2 + |u|^2$$

this is a Hilbert space. Because H is dense in V' and $A+\lambda$ is invertible (for λ large enough, see (2.15)) the space $D(A)$ is dense in V. It follows that V' is a dense subset of the dual space $D(A)'$.

Even smaller and larger spaces are defined by ($k \geq 0$)

$$V_{(2k)} = D(A^k), \qquad V_{(2k+1)} = \{u \in V \mid A^k u \in V\}, \tag{2.57}$$

and their dual spaces defined as

$$V_{(-k)} = V'_{(k)}.$$

We have

$$\ldots \subset V_{(2)} \subset V_{(1)} \subset V_{(0)} \subset V_{(-1)} \subset V_{(-2)} \ldots,$$

where each space is dense in the next space. Note that $V_{(1)} = V, V_{(0)} = H, V_{(-1)} = V'$. If λ sufficiently large (see (2.15)) the operator $A+\lambda$ is an invertible operator $V_{(k+2)} \to V_{(k)}$.

Let $u \in L^2(]0,T[;V_{(k+1)})$, $u' \in L^2(]0,T[;V_{(k)})$, $f \in L^2(]0,T[;V_{(k)})$. The function u is a solution to the PDE if

$$\int_0^T [a(u(t),v(t)) - \langle u'(t), v'(t)\rangle]\,\mathrm{d}t = \int_0^T \langle f(t), v(t)\rangle\,\mathrm{d}t, \tag{2.58}$$

for all v in $C_0^\infty([0,T], V_{(-k+1)})$.

The operator $(A+\lambda)^{\frac{k}{2}}$ is an isomorphism from $V_{(l+k)}$ to $V_{(l)}$ that commutes with the time derivative. It follows that $u \in L^2(]0,T[,V_{(k+1)})$, with $u' \in L^2(]0,T[,V_{(k)})$ is a solution to (2.58) if and only if $(A+\lambda)^{\frac{k}{2}}u$ is a solution in $L^2(]0,T[,V)$ with $((A+\lambda)^{\frac{k}{2}}u)' \in L^2(]0,T[,H)$. Therefore the next result follows immediately from Theorem 2.4.5.

Theorem 2.6.1 *Suppose $a(u,v)$ is a sesquilinear forms on V that obeys (2.15). Suppose $f \in L^2(]0,T[;V_{(k)})$, $u_0 \in V_{(k+1)}, u_1 \in V_{(k)}$. Then there exists a unique solution u of (2.58), (2.19) in $L^2(]0,T[,V_{(k+1)})$, with $u' \in L^2(]0,T[,V_{(k)})$. The solution satisfies in addition*

$$u \in C([0,T],V_{(k+1)}), \qquad u' \in C([0,T],V_{(k)}).$$

It depends continuously on f, u_0, u_1.

Remark 2.6.2 If the coefficients are C^∞ then one can use estimates for pseudodifferential operators (see Hörmander [26], Section 18.1) to show that $V_{(k)} = H_{(k)}(\mathbb{R}^n)$. By integration by parts one can show directly that

$$\langle (A+\lambda)^k u, u\rangle \leq C_1 \|u\|^2_{H_{(k)}}.$$

The sharp Gårding estimate yields that for λ sufficiently large one has

$$\langle (A+\lambda)^k u, u\rangle \geq C_2 \|u\|^2_{H_{(k)}}.$$

Remark 2.6.3 If A is time dependent but such that $D(A(t))$ is independent of t and and isomorphic to $D(A(0))$, then one can still show a result similar to Theorem 2.6.1, see for a result in this direction Lions and Magenes [36] Section 9.5. As in the previous sections the first step is an energy estimate where the energy is now given by

$$E(t) = \tfrac{1}{2}(\langle (A(t)+\lambda)^{k+1}u(t), u(t)\rangle + \langle (A(t)+\lambda)^k u'(t), u'(t)\rangle).$$

One can show that there is a set I_0 that differs from $]0,T[$ by a set of measure zero such that $E(t)$ is Hölder continuous of order $\frac{1}{2}$ on I_0 and it obeys

$$E(t) \leq C\left(E(s) + \int_0^T \|f(\sigma)\|^2_{V_{(k)}}\,\mathrm{d}\sigma\right), \tag{2.59}$$

for all $s, t \in I_0$, $C = C(a, k, T)$. Using this result one can show continuity of solutions and existence and uniqueness.

2.7 Discontinuities at an interface

In this section the theory of Section 2.6 is applied to equation (2.1) in the case that the coefficients have a smooth jump. By this we mean that the medium coefficients $a_{ij;KL}(x), c_{KL}(x)$ of equation (2.1) are discontinuous across a smooth codimension one submanifold, called an interface. We assume that on either side the coefficients are smooth and can be extended smoothly across the interface. We characterize the spaces $V_{(k)}$ defined in the previous section, that contain the solutions and source and initial values. If the source f_K is not too singular at the interface the solutions are in Sobolev spaces on either side and satisfy the well known interface conditions. (Note that a system on both sides of an interface with conditions at the interface can be viewed as a system of double the size with a boundary. For example if the interface is at $x_n = 0$ and x' denotes $(x_1, \dots, x_{n-1})$ this would be the system for the vector $(u_K(x', -x_n), u_K(x', x_n))$ on $x_n > 0$. The interface conditions then become a system of boundary conditions and results like those of Section 2.3 can be obtained.)

We will assume that the interface is at $x_n = 0$, this can be obtained locally by a coordinate transformation. In this section the indices K, L in $a_{ij;KL}, c_{KL}$ will be omitted. However, all arguments will be valid for vector valued u, f and matrix valued coefficients a_{ij}, c. After the transformation of equation (2.56) the operator A is given by

$$-\sum_{i,j} c^{-\frac{1}{2}} \tfrac{\partial}{\partial x_i} a_{ij}(x) \tfrac{\partial}{\partial x_j} c^{-\frac{1}{2}}.$$

and

$$H = L^2(\mathbb{R}^n), \qquad\qquad V = \{u \in L^2(X) \,|\, c^{-\frac{1}{2}}u \in H_{(1)}(\mathbb{R}^n)\}.$$

We will need some function spaces that have not been discussed so far. The anisotropic Sobolev space $H_{(m,s)}$ is defined as the set of all $u \in \mathcal{S}'(\mathbb{R}^n)$ with Fourier transform $\hat{u}(\xi) \in L^2_{\mathrm{loc}}$ and

$$\|u\|^2_{H_{(m,s)}} = (2\pi)^{-n} \int |\hat{u}(\xi)|^2 (1+|\xi|^2)^m (1+|\xi'|^2)^s \,\mathrm{d}\xi < \infty.$$

Thus the elements of $H_{(m,s)}$ have s additional orders of regularity in the x' coordinates.

In this section $\mathbb{R}^n_{x_n \neq 0}$ is viewed as union of two half spaces $\mathbb{R}^n_+ \cup \mathbb{R}^n_-$, where $\mathbb{R}^n_\pm = \{x \in \mathbb{R}^n \mid \pm x_n > 0\}$. Hörmander [26, appendix B] describes the relevant spaces of distributions with a half space as domain. If F is a space of distributions in $\mathbb{R}^n$, then $\bar{F}(\mathbb{R}^n_+)$ is the space of restrictions to $\mathbb{R}^n_+$ of its elements, and $\dot{F}(\overline{\mathbb{R}^n_+})$ is the space of elements supported by $\overline{\mathbb{R}^n_+}$.

For example $\bar{C}^\infty_0(\mathbb{R}^n_+)$ consists of the smooth functions on $\mathbb{R}^n_+$ that can be smoothly extended across the boundary. One can show that these are precisely the elements of $C^\infty(\overline{\mathbb{R}^n_+})$ that vanish outside a compact set. The space $\dot{C}^\infty_0(\overline{\mathbb{R}^n_+})$ is the subspace of functions vanishing of infinite order at $x_n = 0$. The space $\bar{\mathcal{D}}'(\mathbb{R}^n_+)$ (which contains the Sobolev spaces $\bar{H}_{(m,s)}(\mathbb{R}^n_+)$) consists precisely of the linear forms on $\dot{C}^\infty_0(\overline{\mathbb{R}^n_+})$. The space $\dot{\mathcal{D}}'(\overline{\mathbb{R}^n_+})$ consists of the linear forms on $\bar{C}^\infty_0(\mathbb{R}^n_+)$.

For the characterization of $V_{(k)}$, $k \geq 0$ we use the Sobolev space $\bar{H}_{(k)}(\mathbb{R}^n_+ \cup \mathbb{R}^n_-)$. This space can also be viewed as the product $\bar{H}_{(k)}(\mathbb{R}^n_-) \times \bar{H}_{(k)}(\mathbb{R}^n_+)$. If k is a non-negative integer it consists of the L^2 functions on $\mathbb{R}^n$, such that on both sides of the interface derivatives up to order k are in L^2. However, at the interface the function may be discontinuous. The norm is equivalent to

$$\int_{x_n<0} \sum_{|\alpha|\leq k} |\frac{\partial^\alpha u}{\partial x^\alpha}|^2 + \int_{x_n>0} \sum_{|\alpha|\leq k} |\frac{\partial^\alpha u}{\partial x^\alpha}|^2.$$

Although at the interface an element u of $\bar{H}_{(k)}(\mathbb{R}^n_+ \cup \mathbb{R}^n_-)$ may be discontinuous, the restrictions $\frac{\partial^j u}{\partial x_n^j}|_{x_n=0_\pm}$ are well defined in $H_{(k-j-\frac{1}{2})}(\mathbb{R}^{n-1})$ (see Hörmander [26], Theorem B.2.7).

At some point we also use the anisotropic version $\bar{H}_{(k,s)}(\mathbb{R}^n_+ \cup \mathbb{R}^n_-)$. For k a nonnegative integer this space consists of those functions in L^2 such that on each side of the interface $\frac{\partial^j u}{\partial x_n^j}$, $0 \leq j \leq k$ is an L^2 function with values in $H_{(s+k-j)}(\mathbb{R}^{n-1})$. The norm is equivalent to

$$\int_{x_n<0} \sum_{j=0}^{k} \|\frac{\partial^j u}{\partial x_n^j}\|^2_{H_{(s+k-j)}(\mathbb{R}^{n-1})} \,\mathrm{d}x_n + \int_{x_n>0} \sum_{j=0}^{k} \|\frac{\partial^j u}{\partial x_n^j}\|^2_{H_{(s+k-j)}(\mathbb{R}^{n-1})} \,\mathrm{d}x_n.$$

For $V = V_{(1)}$ we find that

$$V = \{u \in \bar{H}_{(1)}(\mathbb{R}^n_+ \cup \mathbb{R}^n_-) \,|\, c^{-\frac{1}{2}} u|_{x_n=0_-} = c^{-\frac{1}{2}} u|_{x_n=0_+}\}.$$

In the following theorem we show that the $V_{(k)}$ are similar. Define the interface differential operators $B_j, j = 0, 1, \dots$ as

$$B_{2k} = c^{-\frac{1}{2}} A^k, \qquad\qquad B_{2k+1} = \sum_j a_{nj} \tfrac{\partial}{\partial x_j} c^{-\frac{1}{2}} A^k.$$

Theorem 2.7.1 *Let the operator A be as given above, let $k \geq 0$. The spaces $V_{(k)}$ defined in (2.57) are given by*

$$V_{(k)} = \{u \in \bar{H}_{(k)}(\mathbb{R}^n_+ \cup \mathbb{R}^n_-) \mid B_j u|_{x_n=0_-} = B_j u|_{x_n=0_+}, j = 0, \dots, k-1\}. \tag{2.60}$$

Proof In this proof we will use the notation $\bar{H}_{(k)} = \bar{H}_{(k)}(\mathbb{R}^n_+ \cup \mathbb{R}^n_-)$. If $u \in \bar{H}_{(k)}$ ($k \geq 2$), and $c^{-\frac{1}{2}} u$ and $\sum_j a_{nj} \frac{\partial}{\partial x_j} c^{-\frac{1}{2}} u$ are continuous at $x_n = 0$, then $Au \in \bar{H}_{(k-2)}$. By induction it follows that $V_{(k)}$ contains the set of $u \in \bar{H}_{(k)}$ that satisfies the conditions in (2.60).

We show that the elements of $V_{(k)}$ satisfy the interface conditions. If $u \in V$ then $c^{-\frac{1}{2}} u \in H_{(1)}$ and hence $c^{-\frac{1}{2}} u$ is continuous at the interface. If in addition $Au \in L^2$ then

$$\sum_{i=1}^{n-1} \sum_j \tfrac{\partial}{\partial x_i} a_{ij} \tfrac{\partial}{\partial x_j} c^{-\frac{1}{2}} u \in L^2(\mathbb{R}, H_{(-1)}(\mathbb{R}^{n-1})). \tag{2.61}$$

Since $\sum_j a_{nj} \frac{\partial}{\partial x_j} c^{-\frac{1}{2}} u \in L^2$ and $\sum_j \frac{\partial}{\partial x_n} a_{nj} \frac{\partial}{\partial x_j} c^{-\frac{1}{2}} u = Au - (2.61)$ it follows that

$$\textstyle\sum_j a_{nj} \frac{\partial}{\partial x_j} c^{-\frac{1}{2}} u \in H_{(1)}(\mathbb{R}, H_{(-1)}(\mathbb{R}^{n-1})),$$

in particular it must be continuous at the interface. If $u \in V_{(2l)}$ then $u, Au, \dots, A^l u \in L^2(x_1)$ and $u, Au, \dots, A^{l-1} u \in V$, hence u in this case satisfies the interface conditions. If $u \in V_{(2l+1)}$ then in addition $A^l u \in V$ and also in this case u satisfies the interface conditions.

It remains to be shown that $V_{(k)}$ is contained in $\bar{H}_{(k)}$. Assume first that $k = 2$. Because $\bar{C}_0^\infty$ is dense in $\bar{H}_{(k)}$ it is sufficient to show that for $u \in \bar{C}_0^\infty$ satisfying the first two interface conditions

$$\|Au\|_{L^2} \geq C_1 \|u\|_{\bar{H}_{(2)}} - C_2 \|u\|_{\bar{H}_{(1)}}. \tag{2.62}$$

Let

$$\begin{aligned} a^{\mathrm{n}}_{ij} &= a_{in} (a_{nn})^{-1} a_{nj}, \\ a^{\mathrm{t}}_{ij} &= a_{ij} - a^{\mathrm{n}}_{ij}, \end{aligned}$$

they represent the ‘normal’ and ‘tangential’ parts of a_{ij} respectively. Note that a^{n} is of the form $(v^t)_i v_j$, the product of a vector v and its transpose v^t. It therefore has a kernel of dimension $N(n-1)$ (recall that we omitted the K, L indices that are in

$\{1,\ldots,N\}$). The tangential part satisfies $a^{\mathrm{t}}_{ij}=0$ if $i=n$ or $j=n$, while the part with $i,j\le n-1$ is positive definite, because it must be positive definite on the kernel of a^{n}.

So we have

$$\|Au\|^2_{L^2}\ge C\Bigg(\int\Big|\sum_{i,j}\tfrac{\partial}{\partial x_i}a^{\mathrm{t}}_{ij}\tfrac{\partial}{\partial x_j}c^{-\frac12}u\Big|^2\,\mathrm{d}x+\int\Big|\sum_{i,j}\tfrac{\partial}{\partial x_i}a^{\mathrm{n}}_{ij}\tfrac{\partial}{\partial x_j}c^{-\frac12}u\Big|^2\,\mathrm{d}x$$
$$+\,2\,\mathrm{Re}\int\sum_{i,j,k,l}(\tfrac{\partial}{\partial x_i}a^{\mathrm{t}}_{ij}\tfrac{\partial}{\partial x_j}c^{-\frac12}u)(\overline{\tfrac{\partial}{\partial x_k}a^{\mathrm{n}}_{kl}\tfrac{\partial}{\partial x_l}c^{-\frac12}u})\,\mathrm{d}x\Bigg). \tag{2.63}$$

Let $\delta^{\mathrm{t}}_{ij}=1$ if $i=j\le n-1$, and 0 otherwise and let $\zeta>0$ be such that $a^{\mathrm{t}}_{ij}-\zeta\delta^{\mathrm{t}}_{ij}$ is still bounded away from zero. The third term can be written as

$$2\,\mathrm{Re}\int\sum_{i,j,k,l}\Big((\zeta\delta^{\mathrm{t}}_{ij}\tfrac{\partial^2}{\partial x_i\partial x_j}c^{-\frac12}u)(\overline{\tfrac{\partial}{\partial x_k}a^{\mathrm{n}}_{kl}\tfrac{\partial}{\partial x_l}c^{-\frac12}u})$$
$$+\,(\tfrac{\partial}{\partial x_i}(a^{\mathrm{t}}_{ij}-\zeta\delta^{\mathrm{t}}_{kl})\tfrac{\partial}{\partial x_j}c^{-\frac12}u)(\overline{\tfrac{\partial}{\partial x_k}a^{\mathrm{n}}_{kl}\tfrac{\partial}{\partial x_l}c^{-\frac12}u})\Big)\,\mathrm{d}x. \tag{2.64}$$

The second part of the integral can be estimated by some constant <1 times the first two terms of (2.63). Indeed the estimate

$$\int\Big|\sum_{i,j}\tfrac{\partial}{\partial x_i}(a^{\mathrm{t}}_{ij}-\zeta\delta^{\mathrm{t}}_{ij})\tfrac{\partial}{\partial x_j}c^{-\frac12}u\Big|^2\,\mathrm{d}x\le(1-\eta)\int\Big|\sum_{i,j}\tfrac{\partial}{\partial x_i}a^{\mathrm{t}}_{ij}\tfrac{\partial}{\partial x_j}c^{-\frac12}u\Big|^2\,\mathrm{d}x$$

(modulo lower order terms) is equivalent to (bringing the left-hand side to the right, and in the integral)

$$\mathrm{Re}\int\sum_{i,j,k,l}(\tfrac{\partial}{\partial x_i}((\sqrt{1-\eta}+1)a^{\mathrm{t}}_{ij}-\zeta\delta^{\mathrm{t}}_{ij})\tfrac{\partial}{\partial x_j}c^{-\frac12}u)$$
$$\times\,(\overline{\tfrac{\partial}{\partial x_k}((\sqrt{1-\eta}-1)a^{\mathrm{t}}_{kl}+\zeta\delta^{\mathrm{t}}_{kl})\tfrac{\partial}{\partial x_l}c^{-\frac12}u})\,\mathrm{d}x\ge0,$$

modulo lower order terms. By partial integration and the sharp Gårding inequality (see Hörmander [26], Theorem 18.1.14) it can be seen that the latter estimate is valid.

To estimate the first term of (2.64) we do two partial integrations, moving $\frac{\partial}{\partial x_k}$ to the left and $\frac{\partial}{\partial x_j}$ to the right. Two boundary terms that arise at the interface due to the partial integration w.r.t. x_n cancel each other (which was the motivation for (2.64)). So the first term of (2.64) equals

$$2\,\mathrm{Re}\int\sum_{i,j,k,l}\zeta\delta_{ij}a^{\mathrm{n}}_{kl}(\tfrac{\partial^2}{\partial x_ix_k}c^{-\frac12}u)(\overline{\tfrac{\partial^2}{\partial x_jx_l}c^{-\frac12}u})\,\mathrm{d}x+\mathrm{l.o.t.} \tag{2.65}$$

The lower order terms involve integrals with less derivatives.

It follows that (with a different constant C)

$$\begin{aligned}\|Au\|_{L^2}^2 \geq C \int \Big[&(\sum_{i,j} \tfrac{\partial}{\partial x_i} a_{ij}^{\mathrm{t}} \tfrac{\partial}{\partial x_j} c^{-\frac{1}{2}} u)^2 + (\sum_{i,j} \tfrac{\partial}{\partial x_i} a_{ij}^{\mathrm{n}} \tfrac{\partial}{\partial x_j} c^{-\frac{1}{2}} u)^2 \\ &+ 2\zeta \sum_{i,j,k,l} \delta_{ij}^{\mathrm{t}} a_{kl}^{\mathrm{n}} (\tfrac{\partial^2}{\partial x_i x_k} c^{-\frac{1}{2}} u)(\tfrac{\partial^2}{\partial x_j x_l} c^{-\frac{1}{2}} u) \Big] \, \mathrm{d}x + \text{l.o.t.}\end{aligned}$$

The result (2.62) follows from this.

For $k > 2$ we use an induction argument. Suppose $V_{(k-2)} \subset \bar{H}_{k-2}$. Then $u \in V_{(k)}$ implies that $u \in \bar{H}_{(k-2)}$ and $Au \in \bar{H}_{(k-2)}$. By Hörmander [26], Theorem B.2.9 it is now sufficient to show that in addition $u \in H_{(2,k-2)}$. So it is sufficient to show that

$$v = (1 - \sum_{i=1}^{n-1} \tfrac{\partial^2}{\partial x_i^2})^{\frac{k-2}{2}} u \in \bar{H}_{(2)}.$$

The function v satisfies the first two interface conditions modulo lower order terms. It follows that in (2.65) lower order boundary terms arise. These can be estimated by lower order Sobolev norms. Therefore estimate (2.62) is valid for v. Therefore $u \in \bar{H}_{(k)}$, and by induction this is true for any $k \geq 0$. □

The result above characterizes the spaces $V_{(k)}$, $k \geq 0$. We now consider the dual spaces $V_{(-k)} = V'_{(k)}$. The spaces $\dot{H}_{(-k)}, \bar{H}_{(-k)}$ are the duals of $\bar{H}_{(k)}, \dot{H}_{(k)}$. We have seen that $V_{(k)}$, $k \geq 0$ is a closed subspace of $\bar{H}_{(k)}(\mathbb{R}_-^n \cup \mathbb{R}_+^n)$. So each linear form on $H_{(k)}$ is also a linear form on $V_{(k)}$. On the other hand it is a consequence of the Hahn-Banach theorem that each linear form $V_{(k)}$ can be extended to a linear form on $\bar{H}_{(k)}$ (see Dunford and Schwartz [18], Theorem II.3.11). So $V_{(-k)}$ is given by $\dot{H}_{(-k)}$ modulo the linear forms in $\dot{H}_{(-k)}$ that vanish on $V_{(k)}$. Note that $V_{(-k)}$ may contain elements that are singular at the interface.

Remark 2.7.2 The text of Elton [19] also discusses hyperbolic differential operators with coefficients that are discontinuous along the plane $x_n = 0$. He constructs spaces H_k^A, where A is a differential operator, that play the same role as the spaces $V_{(k)}$ defined above. These spaces are defined as in (2.60) and it is then shown that the operator $A + \lambda$ is invertible from $V_{(k+2)}$ to $V_{(k)}$. He then shows (among other things) existence and uniqueness results for the solutions of wave equations.

An important difference with our work is that he does not use the variational or divergence form of the operators, where the coefficient is written between the derivatives. The divergence form has the advantage that it allows more general, L^∞ coefficients and that the general theory described in Sections 2.4 and 2.6 can be used. The $V_{(k)}$ are then defined 'canonically' as in (2.60).

Our result Theorem 2.7.1 is also stronger, because Elton did not consider systems of equations and there is an additional assumption on the coefficients, namely that they have well-coupled singularities (see Section 2.4 of [19]).

An element $u \in \dot{H}_{(-k)}$ defines a linear form on $\bar{H}_{(k)}$ and hence in particular it defines a linear form on $\dot{H}_{(k)}$, i.e. an element of $\bar{H}_{(-k)}$, that we denote by $r(u)$. So r defines a map from $\dot{H}_{(-k)}$ to a subset of $\bar{H}_{(k)}$. If $r(u_1) = r(u_2)$ then $u_2 - u_1$ is supported in the interface, $\mathrm{supp}(u_2 - u_1) \subset \partial\mathbb{R}^n_+ \cup \partial\mathbb{R}^n_-$. If follows that $u_2 - u_1$ is a linear form on $(v|_{x_n=\pm 0}, \dots, \frac{\partial^{k-1}v}{\partial x_n^{k-1}}|_{x_n=\pm 0})$. On the other hand, if $u \in \bar{H}_{(-k)}$ is sufficiently regular in the x_n direction at the interface, e.g. in $H_{(0,-k)}$, it defines a linear form on $\bar{H}_{(k)}$ and hence an element $e(u)$ of $\dot{H}_{(-k)}$.

The next theorem shows that in case f is not too singular towards the interface, then the solutions can be obtained by solving the PDE on each side of the interface, and solving interface conditions. We will assume $t \in \mathbb{R}$, to avoid complications due to the initial values. Sometimes the notation $\mathbb{R}_t$ is used to indicate that a function has as variable t.

Theorem 2.7.3 *Let $f \in L^2(\mathbb{R}_t, V_{(-k)})$ and in addition $f \in L^2(\mathbb{R}, \bar{H}_{(-k)}(\mathbb{R}^{n-1} \times \mathbb{R}_t))$ (close to the interface). Then $u \in L^2(\mathbb{R}_t, V_{(-k+1)})$, $u' \in L^2(\mathbb{R}_t, V_{(-k)})$ and*

$$Pu = f, \qquad \textit{weakly} \tag{2.66}$$

if and only if $u = e(\bar{u})$ and $\bar{u} \in L^2(\mathbb{R}_t, \bar{H}_{(-k+1)})$, $\bar{u}' \in L^2(\mathbb{R}_t, \bar{H}_{(-k)})$ satisfies

$$Pu = f, \qquad \textit{weakly on each side of the interface,} \tag{2.67}$$

$$B_i u|_{x_n=0_-} = B_i u|_{x_n=0_+}, \qquad i = 0, 1. \tag{2.68}$$

Proof In this proof the partial differential equation, tested with v is written as

$$\langle u, Av \rangle + \langle u, \frac{\partial^2 v}{\partial t^2} \rangle = \langle f, v \rangle, \tag{2.69}$$

where $v \in C_0^\infty(\mathbb{R}_t, V_{(k+1)})$. For the construction of suitable test functions two sequences of functions $\chi_m(x_n), \psi_m(x_n)$ are needed. Let $\chi_1(x_n) \in C_0^\infty(\mathbb{R})$ be given by

$$\chi_1(x_n) = \begin{cases} 0, & |x_n| > 1, \\ 1, & |x_n| < \frac{1}{4}, \end{cases}$$

and suppose that in addition $\chi_1(x_n) = \chi_1(-x_n)$, $\int \chi_1(x_n)\,\mathrm{d}x_n = 1$. Let $\chi_m(x_n)$ be given by $\chi(mx_n)$. Let $\psi_1(x_n) \in C_0^\infty(\mathbb{R})$ be given by

$$\psi_1(x_n) = \begin{cases} \int_0^{x_n} \chi_1(\sigma)\,\mathrm{d}\sigma, & |x_n| \le 1, \\ 0, & |x_n| \ge 2, \end{cases}$$

and smoothly decreasing on $[-2,-1]$ and $[1,2]$. Let $\psi_m(x_n)$ be given by

$$\psi_m(x_n) = \begin{cases} \frac{1}{m}\psi_1(x - 1 + \frac{1}{m}), & x < -\frac{1}{m}, \\ \frac{1}{m}\psi_1(mx), & -\frac{1}{m} \le x \le \frac{1}{m}, \\ \frac{1}{m}\psi_1(x + 1 - \frac{1}{m}), & x > \frac{1}{m}. \end{cases}$$

Then the second order derivative of $\psi_m(x_n)$ approximates $\delta(x_n - 0_-) - \delta(x_n - 0_+)$ as $m \to \infty$.

We show the 'if' statement. Let $\bar{u} = r(u)$. Clearly (2.66) implies (2.67). It follows from Hörmander [26], Theorem B.2.9 that $\bar{u} \in H_{(2,-k-1)}$. This implies that $B_i u|_{x_n=0_\pm}$ is well defined for $i = 0, 1$.

From the remarks above it follows that $u - e(\bar{u})$ is a distribution with support in $\partial\mathbb{R}^n_- \cup \partial\mathbb{R}^n_+$. This implies that

$$u = e(\bar{u}) + \sum_{j=0,\pm}^{k-2} \langle b_j, \tfrac{\partial^j v}{\partial x_n^j}|_{x_n=0_\pm}\rangle,$$

where the b_j are distributions on $\mathbb{R}^{n-1}$. We show that the b_j vanish. Let $x' = (x_1, \ldots, x_{n-1})$. Define a sequence of test functions v_m by

$$\begin{aligned} v(x', \pm|x_n|, t) &= x_n^k a_{nn}^{-1}(x', 0_\pm) c^{\frac{1}{2}}(x', 0_\pm)\phi(x', t) + O(x_n^{k+1}), \\ v_m(x, t) &= \chi_m(x_n) v(x, t). \end{aligned}$$

By choosing suitable higher order terms the v_m can be made to satisfy the interface conditions up to order k. If $m \to \infty$ all terms in (2.69) go to zero except the term involving b_{k-2} acting on $\frac{\partial}{\partial x_n} a_{nn} \frac{\partial}{\partial x_n} c^{-\frac{1}{2}} v_m$. So (2.69) gives that

$$\langle b_j, \tfrac{\partial}{\partial x_n} a_{nn} \tfrac{\partial}{\partial x_n} c^{-\frac{1}{2}} v_m|_{x_n=0_\pm}\rangle \to 0.$$

Therefore $b_{k-2}(u) = 0$. Choosing v of order x_n^j, $j = k-1, k-2, \ldots, 2$ successively, we find by a similar argument that $b_{k-3}(u), \ldots, b_0(u)$ vanish.

We show that (2.66) implies (2.68). Let $\phi_1 \in C_0^\infty(\mathbb{R}^{n-1} \times \mathbb{R}_t)$ be given. Define

$$v_m(x', \pm|x_n|, t) = a_{nn}^{-1}(x', 0_\pm) c^{\frac{1}{2}}(x', 0_\pm)\phi_1(x', t)\psi_m(x_n) + \chi_m(x_n) w(x, t),$$

where $w(x, t)$ is such that v_m satisfies the interface conditions of order 3 and higher (as far as necessary). Such a w can be found by successively choosing the terms of order $x_n^2, x_n^3, \ldots,$ as far as necessary. It is of order $w = O(x_n^2), x_n \to 0$. Use v_m as a test function in (2.69). Taking the limit $m \to \infty$ we see that the only term that is relevant is the term involving two derivatives w.r.t x_n acting on $\psi_m(x_n)$, the other terms vanish in the limit $m \to \infty$, so we have

$$\int u(x, t)\, c^{-\frac{1}{2}}(x) a_{nn}(x) c^{-\frac{1}{2}}(x) a_{nn}^{-1}(x', 0_\pm) c^{\frac{1}{2}}(x', 0_\pm)\overline{\phi_1(x', t)} \frac{\partial^2 \psi_m(x_n)}{\partial x_n^2}\, \mathrm{d}x\, \mathrm{d}t \to 0.$$

Since $\frac{\partial^2 \psi_m(x_n)}{\partial x_n^2}$ converges to $\delta(x_n - 0_-) - \delta(x_n - 0_+)$ and since u is continuous in x_n (with values in some distribution space) this equation gives (taking the limit $m \to \infty$)

$$\int (c^{-\frac{1}{2}}(x, 0_-) u(x', 0_-, t) - c^{-\frac{1}{2}}(x, 0_+) u(x', 0_+, t))\overline{\phi_1(x', t)}\, \mathrm{d}x'\, \mathrm{d}t = 0.$$

This shows that the first interface condition is satisfied.

By a similar argument equation (2.66) implies the second interface condition. Let $\phi_2(x',t) \in C_0^\infty(\mathbb{R}^{n-1} \times \mathbb{R}_t)$ be arbitrary, and set

$$v_m(x', \pm|x_n|, t) = c^{\frac{1}{2}}(x', 0_\pm)\phi_2(x',t)\chi_m(x_n) + w(x,t)\chi_m(x_n),$$

where $w(x,t)$ is such that v_m satisfies the higher order interface conditions, and w is of order $O(x_n), x_n \to 0$. Taking the limit $m \to \infty$ we obtain

$$\sum_j \langle a_{nj} \tfrac{\partial}{\partial x_j} c^{-\frac{1}{2}} u, c^{\frac{1}{2}} \tfrac{\partial v_m}{\partial x_n} \rangle \to 0.$$

It follows that u satisfies the second interface condition.

We show the 'only if' statement. Suppose v is given that satisfies the interface conditions up to a certain order. We will find $v_{i,m}, i = 1, \ldots, 4$ such that

$$v = v_{1,m} + v_{2,m} + v_{3,m} + v_{4,m},$$

where $v_{4,m}$ is supported in $x_n \neq 0$ and

$$\langle Pu, v_{i,m} \rangle \to 0 \tag{2.70}$$

for $i = 1, 2$ by the interface conditions, and for $i = 3$ because $\|Pv_{3,m}\| \to 0$ in suitable norm, and

$$\langle Pu, v_{4,m} \rangle \to \langle v, f \rangle.$$

To do this let $v_{1,m}$ be given as in the argument for the first interface condition above, with $\phi_1(x',t) = \sum_j a_{nj} \frac{\partial}{\partial x_j} c^{-\frac{1}{2}} v(x', 0, t)$. Let $v_{2,m}$ be as in the second interface condition with $\phi_2(x',t) = c^{-\frac{1}{2}} v(x', 0, t)$. Let $v_{3,m} = \chi_m(v - v_{1,1} - v_{2,1})$ (note that $(v - v_{1,1} - v_{2,1})$ is of order $O(x_n^2)$). Then

$$v_{4,m} = (1 - \chi_m)v + \chi_m v_{1,1} - v_{1,m} + \chi_m v_{2,2} - v_{2,m}.$$

Indeed $v_{4,m}$ is supported in $x_n \neq 0$, and $\|v_{4,m} - v\|_{L^2(\mathbb{R}, H_{(k+1)})} \to 0$. Since $(v - v_{1,1} - v_{2,1})$ is of order $O(x_n^2)$ it follows that $\|Pv_{3,m}\|_{L^2(\mathbb{R}, H_{(k-1)})} \to 0$. By arguments similar to the ones used to prove the 'if' statement the interface conditions imply that (2.70) holds for $i = 1, 2$. This completes the proof. □

2.8 Continuous dependence on the coefficients

It is useful to know how the solutions of wave equations depend on changes of the coefficients, for instance because the coefficients will in general be known only with some finite precision. This is especially true if one wants to determine the coefficients from information about the solutions. In this section certain continuity results will be discussed. We will show that if a certain series of coefficients is convergent in a suitable sense, then the corresponding solutions will converge to the solution of the limiting system.

The coefficients may converge in different sense. We will mainly be interested in the following type of convergence, for the highest order terms

$$\begin{aligned}
&\lim_{m\to\infty} \|(A_m - A)v\|_{V'} \to 0, \qquad \text{for all } v \in V,\\
&\lim_{m\to\infty} |(C_m - C)v| \to 0, \qquad \text{for all } v \in H.
\end{aligned} \tag{2.71}$$

The first result below is also valid when there are lower order terms such that

$$\begin{aligned}
&\lim_{m\to\infty} \|(A_{1,m} - A_1)v\|_{V'} \to 0, \qquad \text{for all } v \in V,\\
&\lim_{m\to\infty} \|(B_m - B)v\|_{V'} \to 0, \qquad \text{for all } v \in H.
\end{aligned} \tag{2.72}$$

For time dependent coefficients we assume the convergence is uniform in t. Below we discuss what this means for the coefficients. The following result is about weak convergence.

Theorem 2.8.1 *Let $a_m(t), a_{1,m}(t), b_m(t), c_m(t)$ and $a(t), a_1(t), b(t), c(t)$ be sesquilinear forms that satisfy the conditions of Lemma 2.5.1 uniformly in m, and such that $a_m(t), a_{1,m}(t), b_m(t), c_m(t)$ converge to $a(t), a_1(t), b(t), c(t)$ in the sense of (2.71), (2.72), uniformly in t. Let $u_m(t)$ be the solutions of equation (2.41) with sesquilinear forms $(a_m, a_{1,m}, b_m, c_m)$ and initial conditions $u_m(0) = u_0 \in V, u'_m(0) = u_1 \in H$. Let $u(t)$ be the solution to (2.41) with initial values u_0, u_1. Then $u_m(t), u'_m(t)$ converge to $u(t), u'(t)$ weakly in $L^2(]0,T[,V)$ and $L^2(]0,T[,H)$ respectively.*

Proof First we note that the solutions u_m exist and are unique by Theorem 2.5.4. The u_m, u'_m are in a bounded subset of $L^2(]0,T[,V)$ and $L^2(]0,T[,H)$ respectively. Since these spaces are reflexive there is a subsequence $m(l)$ such that

$$\begin{aligned}
u_{m(l)} &\to \tilde{u} \qquad \text{weakly in } L^2(]0,T[,V)\\
u'_{m(l)} &\to \chi \qquad \text{weakly in } L^2(]0,T[,H)
\end{aligned}$$

as $l \to \infty$ (cf. [18, II.3.26]). Since the derivative is defined weakly $\chi = \tilde{u}'$.

We will show that $\tilde{u}$ equals u. Choose $v \in C_0^\infty([0,T],V)$. We show that $\tilde{u}$ satisfies the differential equation (2.41). Consider the first term of this equation. We have

$$\begin{aligned}
\langle A_{m(l)}(t)u_{m(l)}(t), v(t)\rangle = {}& \langle A(t)u(t), v(t)\rangle\\
&+ \langle u(t), (A_{m(l)}(t) - A(t))v(t)\rangle + \langle u(t) - u_{m(l)}(t), A(t)v(t)\rangle,
\end{aligned}$$

where we used that A_m and A are symmetric. The second term on the right-hand side converges to 0 by (2.71), the third term because $u_{m(l)} \to u$ weakly. Similar arguments hold for the terms involving A_1, B, C. Therefore $\tilde{u}$ satisfies the differential equation.

The limit $\tilde{u}$ also satisfies the initial conditions by the arguments given in proof of the existence in Theorem 2.4.5.

Because the solution u is unique it follows that not only there is a convergent subsequence, but that in fact $u_m \to u$ weakly. To see this assume that u_m does not converge to u, then there is $\epsilon > 0$, a function v and a subsequence $u_{\tilde{m}(l)}$ such that $\int_0^T \langle u - u_{\tilde{m}(l)}, v\rangle \, \mathrm{d}t > \epsilon$ for all l. However this subsequence must contain a weakly convergent subsequence itself, and by the argument above the limit of this convergent subsequence must be equal to the unique solution u of the system. Contradiction. □

In some cases we can also prove strong convergence. The main step in the proof is to show that the energy of the solutions u_m converges to the energy of the solution u. To prove this we need stronger conditions on the coefficients. We consider the time independent case without lower order terms. There is an additional condition that either $\lambda = 0$ in (2.15) or the embedding of the space V in H is a compact map (mapping bounded sets of V into compact sets in H). If $V = \bar{H}_{(1)}(X)$, or $V = \dot{H}_{(1)}(\overline{X})$, where X is some open of $\mathbb{R}^n$ then this condition is satisfied if X is bounded, this can be shown using Dunford and Schwartz [18], Theorems IV.8.20 and IV.8.21. When $X = \mathbb{R}^n$ (but this is also true for many other unbounded X), we may construct a bounded sequence in V that has no convergent subsequence in H by translating a certain function over a sequence of vectors that goes to infinity.

Theorem 2.8.2 *Assume a_m, c_m and a, c are independent of time and $a_{1,m}, b_m$, and a_1, b are zero. Assume in addition that either the embedding of V in H is compact or that $\lambda = 0$ in (2.15). Suppose that a_m, c_m converge to a, c in the sense of (2.71). Then $u_m \to u$ strongly in $C([0,T], V) \cap C^1([0,T], H)$.*

Proof The energy E_m of the solution satisfies the energy equality (2.48)

$$E_m(t) = \int_0^t \mathrm{Re}\langle f(s), u_m'(s)\rangle \, \mathrm{d}s + \int_0^t \mathrm{Re}\, \lambda \langle u_m(s), u_m'(s)\rangle \, \mathrm{d}s.$$

Because u' converges weakly in $L^2(]0,T[, H)$ the first term on the right-hand side converges to $\int_0^t \langle f(t), u_m'(t)\rangle \, \mathrm{d}t$. If $\lambda = 0$ then the second term is absent. If the embedding of V in H is compact, then because u_m is weakly convergent in $L^2(]0,T[, V)$ it is convergent in $L^2(]0,T[, H)$ and hence the second term converges to $\int_0^t \lambda\langle u(s), u'(s)\rangle \, \mathrm{d}s$. Therefore $E_m(t) \to E(t)$.

Let

$$\xi(t) := a(u - u_m, u - u_m) + \lambda |u - u_m|^2 + c(u' - u_m', u' - u_m').$$

We have

$$\xi(t) = E + E_m - 2a(u, u_m) - 2\lambda\langle u, u_m\rangle - 2c(u', u_m').$$

Because $u_m \to u$ weakly the last three terms converge to $-2E$. Therefore $\xi_m(t) \to 0$, and hence $u_m \to u$ strongly if $m \to \infty$. □

In the next section the question arises whether this is still true when u is less regular. In Section 2.6 we have shown existence and uniqueness of solutions in larger spaces, after transformation of the differential equation. Transforming back to the original variables, gives that if

$$u_0 \in H, \qquad Cu_1 \in V', \qquad f \in L^2(]0,T[,V'),$$

then there exists a unique solution to the initial value problem with

$$u \in C([0,T],H), \qquad Cu' \in C([0,T],V').$$

It turns out that in this case statements similar to the ones above hold. Since the proofs of the these results are very similar to the two proof above we only state the equivalent of Theorem 2.8.2.

Theorem 2.8.3 *Suppose the assumptions of Theorem 2.8.2 are satisfied. Then the solutions u_m of the system A_m, C_m converge to the solution u of the system A, C in the sense that for each t*

$$\begin{aligned} u_m(t) &\to u(t), && \textit{in } H, \\ C_m u_m'(t) &\to Cu'(t), && \textit{in } V'. \end{aligned}$$

Proof From the energy estimates it follows that $u_m, C_m u_m'$ are in a bounded subset of $L^2(]0,T[,H), L^2(]0,T[,V')$, respectively. Since these spaces are reflexive there is a subseries $m(l)$ such that

$$\begin{aligned} u_{m(l)} &\to \tilde{u} && \text{weakly in } L^2(]0,T[,H), \\ C_m u_{m(l)}' &\to \chi && \text{weakly in } L^2(]0,T[,V'), \end{aligned}$$

as $l \to \infty$ (cf. [18, II.3.26]). But $\chi = C\tilde{u}'$.

We show that $\tilde{u}$ satisfies the differential equation. Let v be an arbitrary function in $C_0^\infty([0,T],D(A))$, let $w = (A+\lambda)v$ and let $v_m = (A_m+\lambda)^{-1}w \in C_0^\infty([0,T],D(A_m))$. It follows from the equation

$$A_m - A = (A_m+\lambda) - (A+\lambda) = (A_m+\lambda)[(A+\lambda)^{-1} - (A_m+\lambda)^{-1}](A+\lambda)$$

that

$$\|v_m(t) - v(t)\|_V = \|(A_m+\lambda)^{-1}w(t) - (A+\lambda)^{-1}w(t)\|_V \to 0 \tag{2.73}$$

as $m \to \infty$.

The u_m satisfy the weak differential equation with v_m as test function

$$\int_0^T \Big[\langle A_m u_m, v_m\rangle - \langle C_m u_m', v_m'\rangle - \langle f, v_m\rangle\Big]\,\mathrm{d}t = 0. \tag{2.74}$$

We have

$$\int_0^T \langle f, v_{m(l)}\rangle \,\mathrm{d}t \to \int_0^T \langle f, v\rangle \,\mathrm{d}t,$$

$$\int_0^T \langle C_{m(l)} u'_{m(l)}, v'_{m(l)}\rangle \,\mathrm{d}t \to \int_0^T \langle Cu', v'\rangle \,\mathrm{d}t, \tag{2.75}$$

because $C_{m(l)} u'_{m(l)} \to Cu'$ weakly and (2.73). The first term of (2.74) is equal to

$$\int_0^T \langle A_{m(l)}(t) u_{m(l)}(t), v_{m(l)}(t)\rangle \,\mathrm{d}t = \int_0^T \langle u_{m(l)}(t), w - \lambda(A_{m(l)} + \lambda)^{-1} w\rangle \,\mathrm{d}t.$$

As $l \to \infty$ it converges to

$$\int_0^T \langle u(t), w - \lambda(A + \lambda)^{-1} w\rangle \,\mathrm{d}t = \int_0^T \langle A(t) u(t), v(t)\rangle \,\mathrm{d}t.$$

It follows that

$$\int_0^T \Big[\langle A\tilde{u}, v\rangle - \langle C\tilde{u}', v'\rangle - \langle f, v\rangle\Big] \,\mathrm{d}t = 0,$$

so $\tilde{u}$ satisfies the differential equation.

By arguments similar to the ones in the proof of the existence in Theorem 2.4.5 it follows that $\tilde{u}$ satisfies the correct initial conditions. Using the arguments of Theorem 2.8.1 it follows that there is not only a convergent subsequence, but that in fact $u_m \to u$, weakly.

The energy in this case is

$$E_{(-1),m}(t) = \langle C_m u_m, u_m\rangle + \langle (A_m + \lambda)^{-1} C_m u'_m, C_m u'_m\rangle.$$

It satisfies the following differential equation

$$\frac{dE_{(-1),m}}{dt} = \mathrm{Re}\langle (A + \lambda)^{-1} C_m u'_m, f\rangle + \mathrm{Re}\,\lambda\langle (A + \lambda) C_m u'_m, u_m\rangle.$$

From the transformation (2.56), and equation (2.25) for $(A+\lambda)^{-\frac{1}{2}} u|_{\text{new}}$ it follows that the energy $E_{(1)}$ of a solution u satisfies the following differential equation

$$\frac{dE_{(-1)}}{dt} = \langle (A + \lambda)^{-1} f, Cu'\rangle + \lambda\langle (A + \lambda)^{-1} Cu', Cu\rangle.$$

Using similar arguments as above it follows that $E_{(-1),m}$ converges to $E_{(-1)}(t)$. Together with the weak convergence this implies that

$$\langle C(u - u_m), u - u_m\rangle + \langle (A_m + \lambda)^{-1}(Cu - C_m u'_m), Cu - C_m u'\rangle \to 0.$$

This completes the proof. □

We give sufficient conditions on the coefficients for the operators A_m, C_m to converge in the sense of (2.71) defined above.

Definition 2.8.4 Suppose (E, B, μ) is a measure space, r_m a series of measurable functions on (E, B) then we say $r_m \to 0$ *in μ-measure* if $\lim_{m\to\infty} \mu(E_m^\epsilon) = 0$, for all ϵ, where $E_m^\epsilon = \{x \in E| \, |r_m(x)| > \epsilon\}$.

In case r_m in addition is bounded, it follows by splitting the function r_m as in the second line of (2.77) that this equivalent to $r_m \to 0$ in L^p, for any p. The following lemma states that if $a_{m;ij;KL} \to a_{ij;KL}$, $c_{m;KL} \to c_{KL}$ in measure, then A_m, C_m defined in (2.6) convergence in the sense of (2.71). This lemma shows in particular that the functions $a_{m;ij;KL}, c_{m;KL}$ can be chosen as smooth approximations to $a_{ij;KL}, c_{KL}$.

Lemma 2.8.5 *Let (E, B, μ) be a measure space, let r_m, f_m be measurable functions on (E, B), such that $\|r_m\|_{L^\infty(\mu)} < R$ and $\|f_n\|_{L^2(\mu)} < F$ bounded, and $r_m \to 0$ in μ-measure. Then for each $g \in L^2(\mu)$ we have that*

$$\lim_{m\to\infty} \int_E r_m(x) f_m(x) g(x) \mu(\mathrm{d}x) = 0. \tag{2.76}$$

Proof Suppose (2.76) is not true, then there is an $\eta > 0$ and a subseries $m(l) \nearrow \infty$ such that for all l

$$|\int_E r_{m(l)}(x) f_{m(l)}(x) g(x) \mu(\mathrm{d}x)| \geq \eta.$$

On the other hand we have the estimate

$$\begin{aligned}
&|\int_E r_m(x) f_m(x) g(x) \mu(\mathrm{d}x)| \\
&\quad \leq \int_{E\setminus E_m^\epsilon} |r_m(x)| \, |f_m(x) g(x)| \mu(\mathrm{d}x) + \int_{E_m^\epsilon} |r_m(x)| \, |f_m(x) g(x)| \mu(\mathrm{d}x) \\
&\quad \leq \epsilon \int_E |f_m g| \mu(\mathrm{d}x) + R \int_{E_m^\epsilon} |f_m g| \mu(\mathrm{d}x) \\
&\quad \leq \epsilon F \|g\|_{L^2(\mu)} + RF \left(\int_{E_m^\epsilon} |g(x)|^2 \mu(\mathrm{d}x) \right)^{\frac{1}{2}}
\end{aligned} \tag{2.77}$$

Set $0 < \epsilon < \frac{\eta}{F\|g\|_{L^2(\mu)}}$. The assumption that $r_m \to 0$ in μ-measure implies that there is a subseries $l(k) \nearrow \infty$, for which $\mu(E_{m(l(k))}) \leq 2^{-k}$. This implies that

$$\mu(\bigcup_{k>K} E_{m(l(k))}^\epsilon) \leq 2^{-k} \to 0 \qquad \text{if } k \to \infty.$$

Hence the characteristic functions χ_k of the sets $E_{m(l(k))}^\epsilon$ converge to zero μ-nearly everywhere for $k \to \infty$. Now $\chi_k |g|^2 \leq |g|^2 \in L^1(\mu)$, $\chi_k |g|^2 \to 0$ nearly everywhere, so by Lebesgue's dominated convergence theorem we conclude that

$$\lim_{k\to\infty} \int_{E_{m(l(k))}^\epsilon} |g(x)|^2 \mu(\mathrm{d}x) = 0.$$

It follows that

$$\begin{aligned}\eta &\leq \limsup_{k\to\infty} \left|\int_E r_{m(l(k))} f_{m(l(k))} g\mu(\mathrm{d}x)\right| \\ &\leq \epsilon F\|g\|_{L^2(\mu)} + RF \limsup_{k\to\infty}\left(\int_{E^\epsilon_{m(l(k))}} |g(x)|^2\mu(\mathrm{d}x)\right)^{\frac{1}{2}} \\ &= \epsilon F\|g\|_{L^2(\mu)} < \eta\end{aligned}$$

Contradiction. □

Convergence in the sense above can be compared with other types of convergence. We mention three other types of convergence following Spagnolo [51], Section 4. He discusses the case $V = \dot{H}(X)$, X some open subset of $\mathbb{R}^n$, $N = 1$ (scalar equations) and $c = 1$ (no coefficient in front of the time derivative).

A stronger type of convergence is convergence in operator norm or

$$\|(A_m - A)\phi\|'_V \to 0, \qquad \text{uniform for } \|\phi\|_V \leq 1.$$

This is equivalent to the following equation for the coefficients

$$\|a_{m;ij} - a_{ij}\|_{L^\infty} \to 0,$$

see Spagnolo [51].

A weaker type of convergence is called G-convergence, defined as

$$\|(A_m + \lambda)^{-1} f - (A + \lambda)^{-1} f\|_H \to 0, \qquad \text{for all } f \in V'.$$

Colombini and Spagnolo [15] show that in case the operator A is given by (2.6) and $c = 1$ then $u_m \to u$ weakly in $L^2(]0, T[, \dot{H}(X))$ and strongly in $L^\infty(]0, T[, L^2(S))$ for any compact subset S of X (compare this with the condition that the injection $V \to H$ is compact, this is the case for finite domains). This type of convergence has applications to homogenization.

Weak convergence of the A_m is defined as

$$\langle (A_m - A)\phi, \psi\rangle \to 0, \qquad \text{for all } \phi, \psi \in V.$$

According to Spagnolo [51] there exist sequences A_m that converge weakly and in the sense of G-convergence to different limits, see Example 4, p. 661 in [50]. Therefore this type of convergence is too weak to lead to convergence of solutions.

2.9 Derivative with respect to coefficients

In this section we study the derivative of a solution u of (2.1) with respect to the coefficients. The coefficients $a^{(0)}_{ij;KL}(x), c^{(0)}_{KL}(x)$ are replaced by

$$a^{(1)}_{ij;KL}(X) = a^{(0)}_{ij;KL}(x) + \epsilon\alpha_{ij;KL}(x), \qquad c^{(1)}_{KL}(X) = c^{(0)}_{KL}(x) + \epsilon\gamma_{KL}(x).$$

Let $u^{(j)}$ be the solutions to the initial value problem with coefficients $a^{(j)}_{ij;KL}(x), c^{(j)}_{KL}(x)$, for some source f and initial values u_0, u_1. The Gateaux or directional derivative of the solution with respect to the coefficients is defined as

$$\lim_{\epsilon \to 0} \frac{u^{(1)} - u^{(0)}}{\epsilon}.$$

We show that it can be defined for $a^{(0)}_{ij;KL}(x), c^{(0)}_{KL}(x) \in L^\infty$ and direction (of the derivative) $\alpha_{ij;KL}(x), \gamma_{KL}(x) \in L^\infty$.

Let

$$P^{(j)}_{KL} = -\sum_{i,j} \tfrac{\partial}{\partial x_i} a^{(j)}_{ij;KL}(x) \tfrac{\partial}{\partial x_j} + \tfrac{\partial}{\partial t} c^{(j)}_{KL}(x) \tfrac{\partial}{\partial t},$$

and let V, H be as above (2.6). Let $u^{(j)}$ denote the solution of the partial differential equation with coefficients $a^{(j)}_{ij;KL}(x), c^{(j)}_{KL}(x)$, with initial values $u_0 \in V, u_1 \in H$, and source $f \in L^2(]0,T[,H)$. It follows that

$$P^{(1)}(u^{(1)} - u^{(0)}) = -(P^{(1)} - P^{(0)})u^{(0)}. \tag{2.78}$$

If $\alpha_{ij;KL}(x), \gamma_{KL}(x)$ are bounded then the right hand side of (2.78) is bounded by $C\epsilon$ in $L^2(]0,T[;H_{(-1)})$. Because of energy estimates we have

$$\|u_1 - u_0\|_{L^2(]0,T[,H)} < C\epsilon.$$

If $\epsilon < \epsilon_0$, for ϵ_0 small enough then C depends only on ϵ_0 and P_0, because the energy estimates depend only on bounds for the coefficients. This gives convergence uniform for initial values and source f in some bounded set. However compared to the convergence result of the previous section (which was not uniform) the convergence is weaker by one order.

Dividing equation (2.78) by ϵ gives

$$P^{(1)} \frac{u^{(1)} - u^{(0)}}{\epsilon} = -\frac{P^{(1)} - P^{(0)}}{\epsilon} u_0. \tag{2.79}$$

The formal derivative w.r.t. the coefficients in the direction $\alpha_{ij;KL}(x), \gamma_{KL}(x)$ is given by the value of $\frac{u^{(1)}-u^{(0)}}{\epsilon}$ when $P^{(1)}$ on the left-hand side is replaced by $P^{(0)}$. It is given by the solution v to

$$\begin{aligned} P^{(0)}v &= -\frac{P^{(1)} - P^{(0)}}{\epsilon} u^{(0)}, \\ v(0) &= 0, \qquad\qquad v'(0) = 0. \end{aligned}$$

It follows from Theorem 2.8.3, that (if the assumptions of the theorem are satisfied)

$$\frac{u^{(1)} - u^{(0)}}{\epsilon} \to v, \qquad \text{in } C([0,T],H) \text{ if } \epsilon \to 0. \tag{2.80}$$

Thus in this case the solution has a well defined Gateaux derivative w.r.t. the coefficients.

Remark 2.9.1 Although there is a well defined derivative of the solution w.r.t the coefficients one should note that the convergence in equation (2.80) is in a norm that is one order weaker than the norm of the space containing u. This suggests that, when the coefficients are determined from the solutions using an iterative procedure, one should minimize the difference between real and approximated data in a norm that is different from the norm of the space that contains the data, unlike what one might think naively.

If the initial values u_0, u_1 or the source f are in larger spaces (as in Section 2.6), that is $u_0 \in H$, $Cu_1 \in V'$ or $f \in L^2(]0,T[,V')$, then $u \in C([0,T],H)$ and the right hand side of (2.78) is not in $L^2(]0,T[,D(A_1))$. Hence the argument used above is not valid anymore.

The argument above works for fixed initial values u_0, u_1 and fixed source f. A different question is whether there is some derivative of the map $(u_0, u_1, f) \mapsto u(t)$, in other words whether

$$\frac{u^{(1)} - u^{(0)}}{\epsilon} - v \to 0, \tag{2.81}$$

uniformly for $\|u_0\| \leq 1, |u_1| \leq 1, \|f\|_{L^2(]0,T[,H)} \leq 1$. This is relevant for seismic problems because data is in general taken with a large number of sources.

From equation (2.79) it follows that the left hand side of (2.81) satisfies

$$P^{(1)}\left(\frac{u^{(1)} - u^{(0)}}{\epsilon} - v\right) = -(P^{(1)} - P^{(0)})v.$$

If the coefficients of $P^{(0)}$ and $P^{(1)}$ are in $W^{1,\infty}$ (that is the coefficients themselves and the first order derivatives are in L^∞), then the right hand side is bounded by $C\epsilon$ in $L^2(]0,T[,H_{(-2)}(\mathbb{R}^n))$, and using the results of Section 2.6 it follows that

$$\left\|\frac{u^{(1)} - u^{(0)}}{\epsilon} - v\right\|_{L^\infty(]0,T[,V')} \leq C\epsilon, \tag{2.82}$$

uniformly for $\|u_0\| \leq 1, |u_1| \leq 1, \|f\|_{L^2(]0,T[,H)} \leq 1$. Note however that the convergence is in a norm that is one order weaker.

Seismic data in general involve the distribution kernel of the operator mapping the source to the solutions. An interesting question is whether this distribution kernel can be approximated by linearization in a different norm, for instance a Sobolev norm. Bao and Symes [5] and Bao [4] discuss this is in some detail for the case of an acoustic medium with constant sound speed, where the coefficient of the lower order term is varied.

Chapter 3

Modeling and inversion of seismic data in anisotropic elastic media

Abstract

Seismic data is modeled in the high-frequency approximation, using the techniques of microlocal analysis. We consider general, anisotropic elastic media. Our methods are designed to allow for the formation of caustics. The data is modeled in two ways. First, we give a microlocal treatment of the Kirchhoff approximation — where the medium is assumed to be piecewise smooth, and reflection and transmission occurs at interfaces. Second, we give a refined view on the Born approximation — based upon a linearization of the scattering process in the medium parameters around a smooth background medium. The joint formulation of Born and Kirchhoff scattering allows us to take into account general scatterers as well as the nonlinear dependence of reflection coefficients on the medium parameters. The latter allows the treatment of scattering up to grazing angles.

The outcome of the analysis is a characterization of the singular part of seismic data. We obtain a set of pseudodifferential operators that annihilate the data. In the process we construct a Fourier integral operator, and a reflectivity function, such that the data can be represented by this operator acting on the reflectivity function. In our construction this Fourier integral operator becomes invertible. We give the conditions for invertibility for general acquisition geometry. The result is also of interest for inverse scattering in acoustic media.

3.1 Introduction

In the seismic experiment one generates elastic waves in the Earth using sources at the surface. The waves that return to the surface of the Earth are observed (in fact sources and receivers are not always on the surface of the Earth, this case is also considered). The problem is to reconstruct the elastic properties of the subsurface from the data thus obtained.

The subsurface is given by an open set $X \subset \mathbb{R}^n$. In practice $n = 2$ or 3, but we leave it unspecified. Subsurface position is denoted by x. Sources and receivers are contained in the boundary ∂X of X. Their position is denoted by $\tilde{x}, \hat{x}$. Measurement of data takes places during a time interval $]0, T[$. The set of $(\hat{x}, \tilde{x}, t)$ for which data is taken is called the acquisition manifold Y'; we assume that coordinates y' on Y' are given. We assume that the displacement of the waves is measured for point sources at $\tilde{x}, t = 0$ with all its components, both at the source and at the receiver. Thus we assume that (after preprocessing) the data is given by the Green's function $G_{il}(\hat{x}, \tilde{x}, t)$, for $(\hat{x}, \tilde{x}, t) \in Y'$.

We refer to the codimension of the set of $Y' \in \partial X \times \partial X \times]0, T[$ as the codimension of the acquisition manifold and we denote it by c. For example, in marine data the receivers may lie along a line behind the source, in which case we have $n = 3, c = 1$, $\partial X = \{x \in \mathbb{R}^n \,|\, x_3 = 0\}$, $Y' = \{(\hat{x}, \tilde{x}, t) \in \mathbb{R}^3 \times \mathbb{R}^3 \times]0, T[\,|\, \hat{x}_3 = \tilde{x}_3 = \hat{x}_2 - \tilde{x}_2 = 0\}$. So the data is a function of $2n - 1 - c$ variables. From this data we aim at determining a function of n variables, hence there is a redundancy in the data of dimension $n - 1 - c$.

Our approach follows the work of Beylkin [8, 7], and other authors (see the references below), applying microlocal analysis to the seismic inverse problem. Microlocal analysis and the theory of Fourier integral operators is described in the books by, Hörmander [25, 26, 27], Duistermaat [17], Treves [57, 58].

Beylkin [8] considered the seismic inverse problem in acoustic media with constant density. The data was modeled using the Born approximation, where the scattering is linearized using a linearization in the medium coefficients. The medium perturbation $\delta c(x)$ acts as a distribution of scatterers in a smooth background medium $c(x)$. Given the background medium $c(x)$ an operator was given to reconstruct $\delta c(x)$ microlocally from an n-dimensional subset of the data (from data that is a function of n variables). This was done under certain conditions on the rays. In particular the situation where the wave fronts form caustics was excluded.

When the data is redundant in the sense that the available data is function of more than n variables, then the data can be seen as a family of n-dimensional datasets, where the different n-dimensional subsets in the family are distinguished by part of the coordinates, that we refer to as e (in practical applications this may be the offset coordinate $\hat{x} - \tilde{x}$). For each value of e the inversion can be performed. The result of the inversion, let us call this the reflectivity $r(x, e)$, should not depend on e. This is the criterion that must be used to determine the background medium from the data, see e.g. Symes [55]. Thus under the Assumptions made by Beylkin [8], there is microlocally an invertible map, mapping seismic data to a reflectivity function $r(x, e)$, of which the singular part should not depend on e.

In this chapter we will generalize this result in two directions. First we allow for the presence of caustics. Such a transformation from data to a reflectivity function $r(x, e)$ was previously not defined for data in the neighborhood of a caustic from the scattering point, even in acoustic media. Using Fourier integral operators this restriction is removed. Second we consider general elastic media instead of acoustic media. The propagation of elastic waves is described by a system of partial differential equations. In a constant coefficient medium one can show that there are different modes of propagation, that are independent from each other, i.e. the system decouples. In smoothly varying media this decoupling is no longer exact, but in many cases the system can still be decoupled microlocally, see Taylor [56], Ivrii [30], Dencker [16]. Scattering takes place between the different modes of propagation.

The fact that we consider general elastic media makes the result technically more complicated, and may make it more difficult to see some of the essential ideas, that can also be applied to the acoustic case. On the other hand, there are several good reasons why the results are particularly useful in elastic media. For instance caustics occur much easier in elastic media, they may even occur in elastic media with constant coefficients. Also for elastic media the dependence of reflection coefficients on the scattering angle is more complicated, and it is more important to use this information in the inversion of seismic data.

The data is modeled in two ways. In Section 3.3 we assume that the medium consists of different pieces with smooth interfaces between the different pieces. The medium parameters are assumed to be smooth on each piece, and smoothly extendible across each interface, but they vary discontinuously at the interface. We discuss how to model the high frequency part of the data using Fourier integral operators, following the approach of Taylor [56]. In this way we construct a generalization of the Kirchhoff approximation. In Section 3.4 we discuss the Born approximation. This is essentially a linearization, where the medium parameters are written as the sum of a background medium and a perturbation that is assumed to be small. It is assumed that the background is smooth and that the perturbation contains the singularities of the medium.

The main result is the characterization of seismic data in Theorem 3.7.1. We assume that we have decoupled data for a pair of elastic modes (M, N), where M and N refer to the modes at the receiver and the source respectively. This data can be written as an invertible Fourier integral operator H_{MN} acting on a 'reflectivity' distribution $r_{MN}(x, e)$, that is a function of subsurface position x and an additional variable e, essentially parameterizing the scattering angle and azimuth. The position of the singularities of $r_{MN}(x, e)$ does not depend on e. In the Kirchhoff approximation for elastic media the function $r_{MN}(x, e)$ equals to highest order $R_{MN}(x, e)\|\frac{\partial z_n}{\partial x}\|\delta(z_n(x))$, where $R_{MN}(x, e)$ is the appropriately normalized reflection coefficient for the pair of elastic modes (M, N) and $\|\frac{\partial z_n}{\partial x}\|\delta(z_n(x))$ is the singular function of the interface. For the Born approximation $r_{MN}(x, e)$ is given by pseudodifferential operators that take into account the radiation patterns acting on the medium perturbation.

The result is new even for acoustic media. In that case the coordinate e can be chosen as scattering angle and azimuth. A result in this direction is given in the paper by Xu, Chauris, Lambaré and Noble [59] where such a map is constructed to highest order only for data at acquisition points satisfying a no caustics assumption. They assumed that, given the scattering point, there is a locally diffeomorphic map from the source and receiver coordinates $(\hat{x}, \tilde{x}) \in \partial X \times \partial X$ to the dip and the scattering angle/azimuth (in the notation of Section 3.5 given by $(\xi/\|\xi\|, e)$).

The new step in the proof that is needed to deal with the presence of caustics is given in Section 3.5. The coordinate e is a priori only defined on the coisotropic subset $\mathcal{L}$ of the cotangent acquisition space $T^*Y'\backslash 0$ that contains the wave front set of the data. To construct an invertible Fourier integral operator from data to the function $r_{MN}(x, e)$, the coordinate e has to be defined on an open part of $T^*Y'\backslash 0$. This is done in Lemma 3.5.1, where we construct an extension of the coordinate function e from $\mathcal{L}$ to an open neighborhood of $\mathcal{L}$ in $T^*Y'\backslash 0$. The extension is not unique. Under the no caustics assumption mentioned above there is a 'natural' choice of this extension, which is made implicitly by Xu et al. [59].

The result holds microlocally away from points in the cotangent space $T^*Y'\backslash 0$ that violate our Assumptions 1 to 5, introduced in the main text. The assumptions exclude certain degenerate ray (bicharacteristic) geometries. For example Assumptions 1, 2, 3 exclude rays that go through a singularity of the slowness surface, rays tangent to an interface, and direct rays from source to receiver, respectively. In general, the set of $(y', \eta') \in T^*Y'\backslash 0$ where the assumptions are violated has lower dimension than the dimension of $T^*Y'\backslash 0$. The data associated to such (y', η') can be muted using a pseudodifferential cutoff.

As a consequence of Theorem 3.7.1 we obtain results about the reconstruction of the medium parameters. Given the medium above the interface the function $r_{MN}(x, e)$ and hence the position of the interface and the reflection coefficients can be reconstructed by acting with the inverse H_{MN}^{-1} on the data, see Corollary 3.7.2. For the Born approximation a similar result holds, but an inverse is also obtained directly in Theorem 3.4.5.

When the data is redundant (c sufficiently small) there is in addition a criterion to determine whether the medium above the interface (the background medium in the Born approximation) is correctly chosen. The position of the singularities of the function $r_{MN}(x, e)$, obtained by acting with H_{MN}^{-1} on the data, should not depend on e. There exist pseudodifferential operators $W_{MN}(y', D_{y'})$ that, if the medium above the interface is correctly chosen, annihilate the data, see Corollary 3.7.3. This allows one to do differential semblance optimization [55] in elastic media in the presence of caustics.

We discuss some of the literature on this subject. There have been many publications about high-frequency methods to invert seismic data in acoustic media. The reconstruction of the singular component of the medium coefficients in the Born approximation, without caustics, has been done in the papers by Beylkin [8, 7]. Bleistein [10] discusses the case of a smooth jump using Beylkin's results. It has been shown by Rakesh [46] that the modeling operator in the Born approximation is a Fourier integral operator. Hansen studied the inversion in an acoustic medium with multipathing for both the Born approximation and the case of a smooth jump. Ten Kroode, Smit and Verdel [34] also treat the case of seismic imaging in the presence of multipathing. They discuss in more detail the assumptions (most importantly Assumption 5ii) below) that are made about the geometry of the rays underlying the scattering. Stolk [52] discusses the case when Assumption 5ii) is violated. Nolan and Symes [41] discuss the imaging with different acquisition geometries. The article by Symes [55] discusses the reconstruction of the background medium in the Born approximation.

The mathematical treatment of systems of equations, such as the elastic equations, in the high-frequency approximation has been given by Taylor [56]. This fundamental paper also discusses the interface problem. Beylkin and Burridge [9] discuss the imaging of seismic data in the Born approximation in isotropic elastic media, under a no caustics assumption. De Hoop and Bleistein [23] discuss the imaging in general anisotropic elastic media, using a Kirchhoff-type approximation. The Born approximation in anisotropic elastic media allowing for multipathing is discussed by De Hoop and Brandsberg-Dahl [24].

We give an overview of this chapter. In Section 3.2 we discuss the propagation of waves in smooth, elastic media. First we discuss how asymptotically the elastic system can be decoupled by conjugating with appropriately chosen pseudodifferential operators (a technique that is common in mathematics, but not in the seismic literature). Then we discuss the construction of asymptotic solutions for the decoupled equations using Fourier integral operators. In Section 3.3 we discuss the reflection and transmission of waves at a smooth interface. We explicitly construct Fourier integral operators describing reflected and transmitted waves. These solutions where already discussed, but not explicitly constructed, by Taylor [56]. Thus we prove directly the validity of the Kirchhoff approximation, which is not obvious from e.g. De Hoop and Bleistein [23]. In Section 3.4 we discuss the modeling and inversion of seismic data in the Born approximation. This is important both in its own right and for the reconstruction problem if we model using a smooth jump. We give a comprehensive presentation for the case of general anisotropic media with general acquisition geometry. We discuss in detail the assumptions that are needed. In Section 3.5 we characterize the geometry of the wave front set of the data. Under the assumptions of Section 3.4 this set is contained in a coisotropic submanifold $\mathcal{L}$ of the cotangent space $T^*Y'\backslash 0$. We discuss the extension of symplectic coordinates on $\mathcal{L}$ to a neighborhood of $\mathcal{L}$ in $T^*Y'\backslash 0$. In Section 3.6 we establish microlocally a correspondence between the Kirchhoff approximation and the Born approximation. After the preparations of Sections 3.2 to 3.6 the derivation of our main result in Section 3.7 is relatively simple. We discuss a characterization of seismic data and some consequences, in particular the reconstruction of the position of the interface and the reflection coefficients given the medium above the interface. Finally we construct pseudodifferential operators that annihilate the high-frequency part of the data. In principle these can be used for the reconstruction of the smoothly varying medium parameters above the interface (or of the background medium in the Born approximation).

3.2 Propagation of elastic waves in smoothly varying media

3.2.1 Decoupling the modes

The elastic wave equation is given by

$$\sum_l \left(\rho\, \delta_{il} \frac{\partial^2}{\partial t^2} - \sum_{j,k} \frac{\partial}{\partial x_j} c_{ijkl} \frac{\partial}{\partial x_k} \right) (\text{displacement})_l = (\text{volume force density})_i. \quad (3.1)$$

Here $\rho(x)$ is the volume density of mass and $c_{ijkl}(x)$ is the elastic stiffness tensor, and $i, j, k, l = 1, \ldots, n$.

In order to diagonalize this system, thus decoupling the modes of propagation, it is convenient to remove the x-dependent coefficient ρ in front of the time derivative. Thus we introduce the equivalent system

$$\sum_l P_{il} u_l = f_i, \quad (3.2)$$

where

$$u_l = \sqrt{\rho} (\text{displacement})_l, \qquad f_i = \frac{1}{\sqrt{\rho}} (\text{volume force density})_i, \quad (3.3)$$

and

$$P_{il} = \delta_{il} \frac{\partial^2}{\partial t^2} - \sum_{j,k} \frac{\partial}{\partial x_j} \frac{c_{ijkl}}{\rho} \frac{\partial}{\partial x_k} + \text{l.o.t.} \quad (3.4)$$

is the partial differential operator. Here, we use that ρ is smooth and bounded away from zero. Both systems (3.1) and (3.2) are real, time reversal invariant, and satisfy reciprocity.

We describe how the system (3.2) can be decoupled by transforming it with appropriate pseudodifferential operators, see Taylor [56], Ivrii [30] and Dencker [16]. The goal is to transform the operator P_{il} by conjugation with a matrix-valued pseudodifferential operator $Q(x, D)_{iM}$, $D = D_x = -\mathrm{i}\frac{\partial}{\partial x}$, to an operator that is of diagonal form, modulo a regularizing part,

$$\sum_{i,l} Q(x, D)^{-1}_{Mi} \, P_{il}(x, D, D_t) \, Q(x, D)_{lN} = \operatorname{diag}(P_M(x, D, D_t) \, ; \, M = 1, \ldots, n)_{MN}, \quad (3.5)$$

$D_t = -\mathrm{i}\frac{\partial}{\partial t}$. Here, the indices M, N denote the mode of propagation. In fact for the construction of Fourier integral operator solutions as in the scalar case, it is sufficient to transform the partial differential operator to block-diagonal form, where each of the blocks $P_M(x, D, D_t)$ has scalar principal part (proportional to the identity matrix). In this case we will use the indices M, N to denote the block, and we will omit indices for the components within each block. Let

$$u_M = \sum_i Q(x, D)^{-1}_{Mi} u_i, \qquad f_M = \sum_i Q(x, D)^{-1}_{Mi} f_i. \tag{3.6}$$

The system (3.2) is then equivalent to the uncoupled equations

$$P_M(x, D, D_t) u_M = f_M. \tag{3.7}$$

The time derivative in P_{il} is already on diagonal form, hence we only have to diagonalize its spatial part,

$$A_{il}(x, D) = -\sum_{j,k} \frac{\partial}{\partial x_j} \frac{c_{ijkl}}{\rho} \frac{\partial}{\partial x_k} + \text{l.o.t.}\ .$$

So we have to find Q_{iM} and A_M such that (3.5) is valid with P_{il}, P_M replaced by A_{il}, A_M. The operator P_M is then given by

$$P_M(x, D, D_t) = \frac{\partial^2}{\partial t^2} + A_M(x, D).$$

In view of the properties of stiffness, the principal symbol $A^{\text{prin}}_{il}(x, \xi)$ of $A_{il}(x, D)$ is a positive symmetric matrix, so it can be diagonalized by an orthogonal matrix. On the level of principal symbols, composition of pseudodifferential operators reduces to multiplication. Therefore, we let $Q^{\text{prin}}_{iM}(x, \xi)$ be this orthogonal matrix, and we let $A^{\text{prin}}_M(x, \xi)$ be the eigenvalues of $A^{\text{prin}}_{il}(x, \xi)$, so that

$$\sum_{i,l} Q^{\text{prin}}_{Mi}(x, \xi)^{-1} A^{\text{prin}}_{il}(x, \xi) Q^{\text{prin}}_{lN}(x, \xi) = \operatorname{diag}(A^{\text{prin}}_M(x, \xi))_{MN}. \tag{3.8}$$

The principal symbol $Q^{\text{prin}}_{iM}(x, \xi)$ is the matrix, that has as its columns the orthonormalized polarization vectors associated with the modes of propagation.

If the multiplicities of the eigenvalues $A^{\text{prin}}_M(x, \xi)$ are constant then the principal symbol $Q^{\text{prin}}_{iM}(x, \xi)$ depends smoothly on (x, ξ) and microlocally equation (3.8) carries over to an operator equation. Taylor [56] has shown that if this condition is satisfied then decoupling can be accomplished to all orders, where each block corresponds to a different eigenvalue. In fact he proved the following slightly more general result.

Lemma 3.2.1 *(Taylor) Suppose the pseudodifferential operator $Q_{iM}(x, D)$ of order 0 is such that*

$$\sum_{i,l} Q(x, D)^{-1}_{Mi} A(x, D)_{il} Q(x, D)_{lN} = \begin{pmatrix} A_{(1)}(x, D) & 0 \\ 0 & A_{(2)}(x, D) \end{pmatrix}_{MN} + a(x, D)_{MN},$$

where the symbols $A_{(1)}(x,\xi), A_{(2)}(x,\xi)$ are homogeneous of order 2, and $a(x,\xi)_{MN}$ is polyhomogeneous of order 1. Suppose the spectra of $A_{(1)}(x,\xi), A_{(2)}(x,\xi)$ are disjoint on a conic neighborhood of some $(x_0,\xi_0) \in T^*X\backslash 0$. Then by modifying Q with lower order terms the system can be transformed such that

$$a(x,D)_{MN} = \begin{pmatrix} a_{(1)}(x,D) & 0 \\ 0 & a_{(2)}(x,D) \end{pmatrix}_{MN} + \text{smoothing remainder},$$

microlocally around (x_0,ξ_0).

This implies that if the multiplicity of a particular eigenvalue $A_M^{\text{prin}}(x,\xi)$ is constant, then the system can be transformed such that the part related to this eigenvalues decouples from the rest of the system, modulo a smoothing remainder. In this work we will assume that at least some of the modes decouple (microlocally). This is stated as Assumption 1 below. At that point we will also discuss whether this assumption is satisfied in relevant cases.

We now give an alternative characterization of the quantities $A_M^{\text{prin}}(x,\xi)$ and $Q_{iM}^{\text{prin}}(x,\xi)$. The values $\tau = \pm\sqrt{A_M^{\text{prin}}(x,\xi)}$ are precisely the solutions to the equation

$$\det P_{il}^{\text{prin}}(x,\xi,\tau) = 0. \tag{3.9}$$

The multiplicity of $A_M^{\text{prin}}(x,\xi)$ is equal to the multiplicity of the corresponding root of (3.9). The columns of $Q_{iM}^{\text{prin}}(x,\xi)$ satisfy

$$Q_{iM}^{\text{prin}} \in \ker P_{il}^{\text{prin}}(x,\xi,\sqrt{A_M^{\text{prin}}(x,\xi)}).$$

Since $P_{il}^{\text{prin}}(x,\xi,\tau)$ is homogeneous in (ξ,τ), one may choose to use the slowness vector $-\tau^{-1}\xi$ instead of the cotangent or wave vector ξ in calculations. The set of $-\tau^{-1}\xi$ such that (3.9) holds is called the slowness surface, which can be easily visualized. A section of the slowness surface for the case of a transversely isotropic medium in 3 dimensions is given in Figure 3.1a. Note that the slowness surface need not be convex. The multiplicity of the eigenvalues changes at the points (directions) were the different sheets intersect.

The second-order equations (3.7) clearly describe the decoupling of the original system into different elastic modes. In addition, equations (3.7) inherit the symmetries of the original system. It is easy to see that they are time reversal invariant. The operators $Q_{iM}(x,D), A_M(x,D)$ can be chosen in such a way that $Q_{iM}(x,\xi) = -\overline{Q_{iM}(x,-\xi)}$, $A_M(x,\xi) = \overline{A_M(x,\xi)}$. Then Q_{iM}, A_M are real. We argue that equations (3.7) also satisfy reciprocity. For the causal Green's function $G_{ij}(x,x_0,t-t_0)$ reciprocity means that $G_{ij}(x,x_0,t-t_0) = G_{ji}(x_0,x,t-t_0)$. We show that such a relationship also holds (modulo smoothing operators) for the Green's function $G_M(x,x_0,t-t_0)$ associated with (3.7). The transpose operator $Q(x,D)^t_{Mi}$ (obtained by interchanging x,x_0 and i,M in the distribution kernel $\mathcal{Q}_{iM}(x,x_0)$ of

$Q_{iM}(x,D))$ is also a pseudodifferential operator, with principal symbol $Q^{\text{prin}}(x,\xi)^t_{Mi}$. As noted before for the principal symbol, it follows from the fact that $A^t_{ij} = A_{ij}$ that we can choose Q orthogonal, i.e. such that $\sum_M Q(x,D)_{iM} Q(x,D)^t_{Mj} = \delta_{ij}$. From the fact that

$$G_M(x,x_0,t-t_0) = Q(x,D)^{-1}_{Mi} G_{ij}(x,x_0,t-t_0) Q(x_0,D_{x_0})_{jM}$$

it then follows that microlocally G_M is reciprocal, i.e. $G_M(x,x_0,t-t_0) = G_M(x_0,x,t-t_0)$, modulo smoothing operators.

Remark 3.2.2 We already observed that if an eigenvalue $A^{\text{prin}}_M(x,\xi)$ has constant multiplicity $m_M > 1$, then u_M is an m-dimensional vector and (3.7) is a $m_M \times m_M$ system, with scalar principal symbol. For such a system a microlocal solution can be constructed in the same way as for scalar systems, see the next subsection. In this case all kinematic quantities, such as bicharacteristics, phase functions, canonical relations depend only on M. Other quantities such as u_M and $Q_{iM}(x,D)$ will have multiple components. The Green's function G_M and its amplitude $\mathcal{A}_M$, to be introduced above (3.20), then are $m_M \times m_M$ matrices. To simplify notation we do not take this into account explicitly. However, the reader can check that the results of this work can be generalized to this case.

3.2.2 The Green's function

To evaluate the Green's function we use the first-order system for u_M that is equivalent to (3.7). It is given by

$$\frac{\partial}{\partial t}\begin{pmatrix} u_M \\ \frac{\partial u_M}{\partial t} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -A_M(x,D) & 0 \end{pmatrix} \begin{pmatrix} u_M \\ \frac{\partial u_M}{\partial t} \end{pmatrix} + \begin{pmatrix} 0 \\ f_M \end{pmatrix}. \tag{3.10}$$

This system can be decoupled also. Let $B_M(x,D) = \sqrt{A_M(x,D)}$, which is a pseudodifferential operator of order 1 that exists because $A_M(x,D)$ is positive definite. The principal symbol of $B_M(x,D)$ is given by $B^{\text{prin}}_M(x,\xi) = \sqrt{A^{\text{prin}}_M(x,\xi)}$. We find that then (3.10) is equivalent to the following two first-order equations

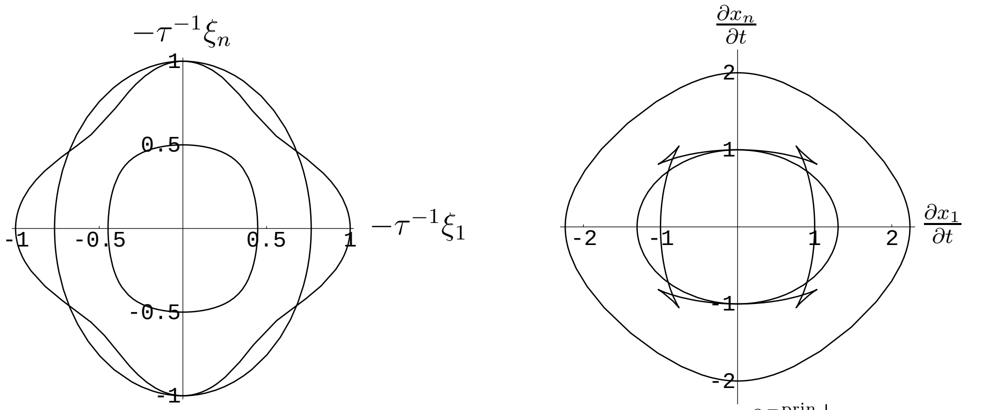
$$\left(\frac{\partial}{\partial t} \pm \mathrm{i} B_M(x,D)\right) u_{M,\pm} = f_{M,\pm}, \tag{3.11}$$

upon transforming

$$\begin{aligned} u_{M,\pm} &= \tfrac{1}{2} u_M \pm \tfrac{1}{2}\mathrm{i} B_M(x,D)^{-1} \frac{\partial u_M}{\partial t}, \\ f_{M,\pm} &= \pm \tfrac{1}{2}\mathrm{i} B_M(x,D)^{-1} f_M. \end{aligned} \tag{3.12}$$

We construct operators $G_{M,\pm}$ with distribution kernel $G_{M,\pm}(x,x_0,t)$ that solve the initial value problem for (3.11). Then using Duhamel's principle we find that

$$u_{M,\pm}(x,t) = \int_0^t G_{M,\pm}(x,x_0,t-t_0) f_{M,\pm}(x_0,t_0)\,\mathrm{d}x_0\,\mathrm{d}t_0.$$



(a) Cotangent: $B_M^{\text{prin}}(x,\tau^{-1}\xi) = 1$ (b) Tangent: $\left.\frac{\partial B_M^{\text{prin}}}{\partial \xi}\right|_{B_M^{\text{prin}}(x,\tau^{-1}\xi)=1}$

Figure 3.1: (a) Section of a slowness surface (the characteristic surface) for a transversely isotropic medium in $n = 3$ dimensions. (b) Set of velocities associated to the slowness surface in a). Note the caustics that occur due to the fact that one of the sheets is not convex.

It follows from (3.12) that the Green's function for the second-order decoupled equation is then given by

$$G_M(x, x_0, t) = \tfrac{1}{2}\mathrm{i}G_{M,+}(x, x_0, t)B_M(x_0, D_{x_0})^{-1} - \tfrac{1}{2}\mathrm{i}G_{M,-}(x, x_0, t)B_M(x_0, D_{x_0})^{-1}. \tag{3.13}$$

The operators $G_{M,\pm}$ are Fourier integral operators. Their construction is well known, see e.g. Duistermaat [17], Chapter 5. The singularities are propagated along the bicharacteristics, that are determined by Hamilton's equations generated by the principal symbol (factor i divided out) $\tau \pm B_M^{\text{prin}}(x, \xi)$ of (3.11). These equations read

$$\begin{aligned} \frac{\partial x}{\partial \lambda} &= \pm\frac{\partial}{\partial \xi}B_M^{\text{prin}}(x,\xi), & \frac{\partial t}{\partial \lambda} &= 1, \\ \frac{\partial \xi}{\partial \lambda} &= \mp\frac{\partial}{\partial x}B_M^{\text{prin}}(x,\xi), & \frac{\partial \tau}{\partial \lambda} &= 0. \end{aligned} \tag{3.14}$$

The solution may be parameterized by t. We denote the solution of (3.14) with the $+$ sign and initial values x_0, ξ_0 by $(x_M(x_0, \xi_0, t), \xi_M(x_0, \xi_0, t))$. The solution with the $-$ sign is found upon reversing the time direction, in other words, it is given by $(x_M(x_0, \xi_0, -t), \xi_M(x_0, \xi_0, -t))$. Observe that the group velocity (the velocity $\frac{\partial x}{\partial t}$ of the bicharacteristic) is orthogonal to the slowness surface. Where the slowness surface fails to be convex, caustics may arise instantly from a point source. An example is shown in Figure 3.1b.

The canonical relation of the operator $G_{M,\pm}$ is given by

$$C_{M,\pm} = \{(x_M(x_0, \xi_0, \pm t), t, \xi_M(x_0, \xi_0, \pm t), \mp B_{M,\pm}(x_0, \xi_0); x_0, \xi_0)\}. \tag{3.15}$$

A convenient choice of phase function is described in Maslov and Fedoriuk [38]. They state that one can always use a subset of the cotangent vector components as phase variables. Let us choose coordinates for $C_{M,+}$ of the form

$$(x_I, x_0, \xi_J, \tau), \tag{3.16}$$

where $I \cup J$ is a partition of $\{1, \ldots, n\}$. It follows from Theorem 4.21 in Maslov and Fedoriuk [38] that there is a function $S_{M,+}(x_I, x_0, \xi_J, \tau)$, such that locally $C_{M,+}$ is given by

$$\begin{aligned} x_J &= -\frac{\partial S_{M,+}}{\partial \xi_J}, & t &= -\frac{\partial S_{M,+}}{\partial \tau}, \\ \xi_I &= \frac{\partial S_{M,+}}{\partial x_I}, & \xi_0 &= -\frac{\partial S_{M,+}}{\partial x_0}. \end{aligned} \tag{3.17}$$

Here we take into account the fact that $C_{M,+}$ is a canonical relation, which introduces a minus sign for ξ_0. A nondegenerate phase function for $C_{M,+}$ is then found to be

$$\phi_{M,+}(x, x_0, t, \xi_J, \tau) = S_{M,+}(x_I, x_0, \xi_J, \tau) + \langle \xi_J, x_J \rangle + \tau t. \tag{3.18}$$

On the other hand, the canonical relation $C_{M,-}$ is given by

$$C_{M,-} = \{(x, t, -\xi, -\tau; x_0, -\xi_0) \mid (x, t, \xi, \tau; x_0, \xi_0) \in C_{M,+}\}.$$

Thus a phase function for $C_{M,-}$ is $\phi_{M,-}(x, x_0, t, \xi_J, \tau) = -\phi_{M,+}(x, x_0, t, -\xi_J, -\tau)$. We may define the canonical relation for G_M as $C_M = C_{M,+} \cup C_{M,-}$ and a phase function $\phi_M = \phi_{M,-}$ if $\tau > 0$, $\phi_M = \phi_{M,+}$ if $\tau < 0$.

We have to assume that the decoupling is valid microlocally around the bicharacteristic. In that case Theorem 5.1.2 of Duistermaat [17] implies that the operator $G_{M,\pm}$ is microlocally a Fourier integral operator of order $-\frac{1}{4}$. Hence, microlocally we have an expression for $G_{M,\pm}$ of the form

$$G_{M,\pm}(x, x_0, t) = (2\pi)^{-\frac{|J|+1}{2} - \frac{2n+1}{4}} \int \mathcal{A}_{M,\pm}(x_I, x_0, \xi_J, \tau) \mathrm{e}^{\mathrm{i}\phi_{M,\pm}(x, x_0, t, \xi_J, \tau)} \, \mathrm{d}\xi_J \, \mathrm{d}\tau. \tag{3.19}$$

The factors of (2π) in front of the integral are according to the convention of Treves [58] and Hörmander [27].

The amplitude $\mathcal{A}_{M,\pm}(x_I, x_0, \xi_J, \tau)$ satisfies a transport equation along the bicharacteristics $(x_M(x_0, \xi_0, \pm t), \xi_M(x_0, \xi_0, \pm t))$. Properties of amplitudes are described for instance in Treves [58], Section 8.4. The amplitude is an element of $M_{C_M} \otimes \Omega^{1/2}(C_M)$, the tensor product of the Keller-Maslov bundle M_{C_M} and the half-densities on the canonical relation C_M. If the subprincipal part of $A_M(x, D)$ is a matrix, then the amplitude is also a matrix, see Remark 3.2.2. The Keller-Maslov bundle gives a factor i^k, where k is an index, which we will absorb in the amplitude. So the amplitude should be seen as a function on the canonical relation $C_{M,\pm}$, coordinatized by (x_I, x_0, ξ_J, τ), see (3.16). It is possible to choose a Maslov phase function with a different set of phase variables, for instance $\xi_{\tilde{J}}$ (and not τ), where $(\tilde{I}, \tilde{J})$ is a partition of $\{1, \dots, n\}$ and $C_{M,\pm}$ is parameterized by $(x_{\tilde{I}}, x_0, t, \xi_{\tilde{J}})$. In that case the transformed amplitude $\tilde{\mathcal{A}}_{M,\pm}(x_{\tilde{I}}, x_0, t, \xi_{\tilde{J}})$ contains a Jacobian factor to the power one half, i.e.

$$\left|\tilde{\mathcal{A}}_{M,\pm}(x_{\tilde{I}}, x_0, t, \xi_{\tilde{J}})\right| = \left|\mathcal{A}_{M,\pm}(x_I, x_0, \xi_J, \tau)\right| \left|\frac{\partial(x_I, x_0, \xi_J, \tau)}{\partial(x_{\tilde{I}}, x_0, t, \xi_{\tilde{J}})}\right|^{\frac{1}{2}}, \tag{3.20}$$

where in the Jacobian both sets of variables are coordinates on $C_{M,\pm}$.

We will calculate the left-hand side of (3.20). For this purpose, consider the Green's function $G_{M,\pm}(x, x_0, t - t_0)$ with t and $t_0 = 0$ fixed. It can be viewed as an invertible Fourier integral operator, mapping the displacement at $t = 0$, $u|_{t=0} \in \mathcal{E}'(X)$ to the displacement at t, $u|_t \in \mathcal{D}'(X)$, with phase $\tilde{\phi}_{M,\pm}(x, x_0, t, \xi_{\tilde{J}})$ and amplitude $\tilde{\mathcal{A}}_{M,\pm}(x_{\tilde{I}}, x_0, t, \xi_{\tilde{J}})$. To highest order the energy at time t is given by

$$\int \left|B_M(x, D) u_{M,\pm}(x, t)\right|^2 \,\mathrm{d}x.$$

Conservation of this quantity is reflected by the relation

$$G_{M,\pm}(t)^* B_M(x, D)^* B_M(x, D) G_{M,\pm}(t) = B_{M,\pm}(x_0, D_{x_0})^* B_{M,\pm}(x_0, D_{x_0}),$$

where the left-hand side denotes a composition of Fourier integral operators and * denotes the adjoint. Since the left-hand side is a product of invertible Fourier integral operators, we can use the theory of Section 8.6 in Treves [58]. We find that to highest order

$$\left|(2\pi)^{-\frac{1}{4}} \tilde{\mathcal{A}}_{M,\pm}(x_{\tilde{I}}, x_0, t, \xi_{\tilde{J}})\right|^2 = \left|\det \frac{\partial \xi_0}{\partial(x_{\tilde{I}}, \xi_{\tilde{J}})}\right| \left|\frac{B_M(x_0, \xi_0)}{B_M(x, \xi)}\right|^2.$$

The value of $B_M(x, \xi)$ equals the frequency τ and is conserved along the bicharacteristic. Recall that (x_0, ξ_0, t) are valid coordinates for $C_{M,\pm}$ (cf. (3.15)). The Jacobian $\left|\frac{\partial(x_0,\xi_0,t)}{\partial(x_I,x_0,t,\xi_J)}\right|$ is equal to the factor $\left|\det \frac{\partial \xi_0}{\partial(x_I,\xi_J)}\right|$. It follows that to highest order

$$\left|\tilde{\mathcal{A}}_{M,\pm}(x_{\tilde{I}}, x_0, t, \xi_{\tilde{J}})\right| = (2\pi)^{\frac{1}{4}} \left|\det \frac{\partial(x_0, \xi_0, t)}{\partial(x_{\tilde{I}}, x_0, t, \xi_{\tilde{J}})}\right|^{\frac{1}{2}}. \tag{3.21}$$

From (3.20) it now follows that

$$\left|\mathcal{A}_{M,\pm}(x_I, x_0, \xi_J, \tau)\right| = (2\pi)^{\frac{1}{4}} \left|\det \frac{\partial(x_0, \xi_0, t)}{\partial(x_I, x_0, \xi_J, \tau)}\right|^{\frac{1}{2}}. \tag{3.22}$$

We give our result about the Green's function for (3.7), collecting the results of this section, and using equations (3.12) and (3.22), to obtain a statement about the amplitude. We will assume that microlocally around the relevant bicharacteristics the decoupling is valid. Let $\mathrm{Char}(P_M)$ be the characteristic set of P_M given by $\{(x, t, \xi, \tau) \mid P(x, \xi, \tau) = 0\}$. The Green's function is such that precisely the singularities of f_M at $\mathrm{Char}(P_M)$ propagate (see Hörmander [26], Theorem 23.2.9). Thus we have

Assumption 1 *On a neighborhood of the bicharacteristic the multiplicity of the eigenvalue $A_M^{\mathrm{prin}}(x,\xi)$ in (3.8) is constant.*

Lemma 3.2.3 *Suppose that for the bicharacteristics through $\mathrm{WF}(f_M) \cap \mathrm{Char}(P_M)$ Assumption 1 is satisfied. Then u_M is given microlocally, away from $\mathrm{WF}(f_M)$, by*

$$u_M(x,t) = \int G_M(x,x_0,t-t_0) f_M(x_0,t_0)\,\mathrm{d}x_0\,\mathrm{d}t_0, \tag{3.23}$$

where $G_M(x,x_0,t)$ is the kernel of a Fourier integral operator with canonical relation C_M and order $-1\frac{1}{4}$, mapping functions of x_0 to functions of (x,t). It can be written as

$$G_M(x,x_0,t) = (2\pi)^{-\frac{|J|+1}{2}-\frac{2n+1}{4}} \int \mathcal{A}_M(x_I,x_0,\xi_J,\tau) \mathrm{e}^{\mathrm{i}\phi_M(x,x_0,t,\xi_J,\tau)}\,\mathrm{d}\xi_J\,\mathrm{d}\tau. \tag{3.24}$$

For the amplitude $\mathcal{A}_M(x_I,x_0,\xi_J,\tau)$ we have to highest order

$$|\mathcal{A}_M(x_I,x_0,\xi_J,\tau)| = (2\pi)^{\frac{1}{4}} \tfrac{1}{2} |\tau|^{-1} \left| \det \frac{\partial(x_0,\xi_0,t)}{\partial(x_I,x_0,\xi_J,\tau)} \right|^{\frac{1}{2}}. \tag{3.25}$$

The implications of Assumption 1 for elastic media depend on which class of media one is interested in. By a class of media we mean a set of media parameterized by a number of parameters. From a physical point of view one may be interested in media where the elastic tensor is characterized by certain symmetry properties.

Isotropic media are characterized by the mass density ρ and the Lamé parameters λ and μ. The matrix $A_{il}^{\mathrm{prin}}(x,\xi)$ has two eigenvalues, $A_1^{\mathrm{prin}}(x,\xi) = \frac{\lambda+2\mu}{\rho}\|\xi\|^2$ with polarization vector proportional to ξ (referred to as the P-mode), and $A_2^{\mathrm{prin}}(x,\xi) = \frac{\mu}{\rho}\|\xi\|^2$ with polarization space normal to ξ (the two S-modes). Thus this system can be decoupled.

If the matrix $A_{il}^{\mathrm{prin}}(x,\xi)$ of an isotropic medium is perturbed by a small amount, then one eigenvalue of the perturbed matrix will be close to the P-eigenvalue of the isotropic medium, and two eigenvalues will be close to the S-eigenvalue. The two eigenvalues close to the S-eigenvalue of the isotropic medium will not coincide in general, but may coincide for certain values of (x,ξ). So in elastic media sufficiently close to an isotropic medium there will still be a quasi-P mode that decouples from the other modes, but the two quasi-S modes will in general not decouple.

The elastic system for generic elastic media has been investigated by Braam and Duistermaat [12]. The set of singular points is generically of codimension three (thus one lower than one would expect naively), and is of conical form in the neighborhood of the singular point. They give a normal form for such systems and investigate the behavior of its associated bicharacteristics and polarization spaces. In this case the system cannot be decoupled. However, in a generic elastic medium there cannot be an open set of bicharacteristics that pass through a singular point, since the singular points form a set of codimension 3. In this sense the set of bicharacteristics that have to be excluded is small.

In case the elastic tensor has symmetries it is determined by less than 21 coefficients. The characteristic sets of such media are analyzed by Musgrave [40]. In this case the singularities can be of different types. For instance, in some classes of media, such as transversely isotropic media, the determinant factors into smooth factors. In that case the multiplicities of the eigenvalues $A_M^{\text{prin}}(x, \xi)$ can vary on a larger (codimension 2) subset of $T^*X\backslash 0$. Since the bicharacteristics are curves on a codimension 1 surface, Assumption 1 can be violated on an open set of bicharacteristics.

3.3 Reflection at an interface: Microlocal analysis of the 'Kirchhoff' approximation

A particular way to model the subsurface is to assume that it consists of different layers that have different physical properties, in our case the elastic coefficients c_{ijkl} and the density ρ. In this section, we will model the reflection of waves at a smooth interface between two such layers with smoothly varying medium parameters.

The amplitude of the scattered waves is determined essentially by the reflection coefficients, and implicitly by the curvature of the interface. It is well known how to calculate these for two constant coefficient media and a plane interface (see e.g. Aki and Richards [2], Chapter 5). In the case of smoothly varying media they determine the scattering in the high-frequency limit, see Taylor [56] for a treatment of reflection and transmission of waves using microlocal analysis. For the acoustic case, see also Hansen [22].

Mathematically the reflection and transmission of waves is formulated as a boundary value problem. The displacement u_l must satisfy the partial differential equation and initial conditions. In addition the displacement and the normal traction have to be continuous at the interface. Denote by ν the normal to the interface. The following equations must hold

$$\begin{aligned} \sum_l P_{il} u_l &= f_i && \text{away from the interface,} \\ u_l &= 0 && \text{for } t < 0, \end{aligned} \tag{3.26}$$

while

$$\begin{aligned} &\rho^{-1/2} u_l && \text{is continuous at the interface,} \\ &\sum_{j,k,l} \nu_j c_{ijkl} \frac{\partial}{\partial x_k} (\rho^{-1/2} u_l) && \text{is continuous at the interface.} \end{aligned} \tag{3.27}$$

Here, we have the factors ρ because of our normalization (3.3). We assume that the source vanishes on a neighborhood of the interface. That this is a well-posed problem can for instance be shown using energy estimates (see e.g. Lions and Magenes [36], Section 3.8).

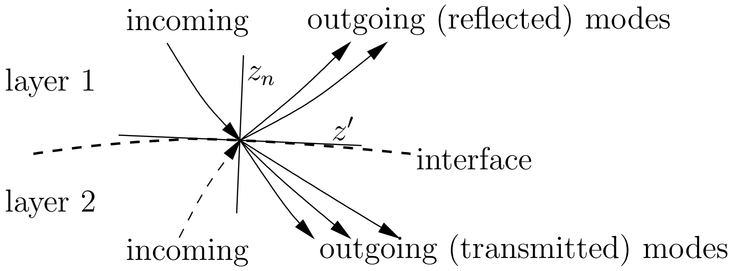


Figure 3.2: Incoming and outgoing rays

The solutions to the partial differential equation with $f = 0$ follow from the theory discussed in Section 3.2. The singularities are propagated along the bicharacteristics, curves in $T^*(X \times \mathbb{R})\backslash 0$, given by

$$(x_M(x_0, \xi_0, \pm t), t, \xi_M(x_0, \xi_0, \pm t), \mp B_M(x_0, \xi_0)).$$

This is the bicharacteristic associated with the $M, \pm$ constituent of the solution, see Section 3.2. We define a bicharacteristic to be incoming if its direction is from inside a layer towards the interface for increasing time. We define a bicharacteristic to be outgoing if its direction is away from the interface into a layer for increasing time, see Figure 3.2.

Assume that the incoming bicharacteristic stays inside a layer from $t = 0$ until it hits the interface, then the solution along such a bicharacteristic is determined completely by the partial differential equation and the initial condition. On the other hand, the solution along the outgoing bicharacteristics is not determined by the partial differential equation and the initial condition: We will show that the solution along the outgoing bicharacteristics is determined by the partial differential equation and the interface conditions in (3.27).

Let us consider the consequences of the interface conditions. Assume for the moment that the interface is located at $x_n = 0$. We denote $x' = (x_1, \ldots, x_{n-1}), x = (x', x_n)$ and similarly for ξ. The wavefront set of the restriction of u_l to $x_n = 0$ satisfies

$$\mathrm{WF}(u_l|_{x_n=0}) = \{(x', t, \xi', \tau) \,|\, \text{there is } \xi_n \text{ with } (x', 0, t, \xi', \xi_n, \tau) \in \mathrm{WF}(u_l)\}.$$

It follows that a solution traveling along a bicharacteristic that intersects the boundary at some point $(x', 0, t)$ interacts with any other such solution as long as the associated values for ξ', τ in their wavefront sets coincide (Snell's law). This is depicted in Figure 3.3.

Depending on the boundary coordinate x' and the 'tangential' slowness $-\tau^{-1}\xi'$, the number of interacting bicharacteristics may vary. For large values of $-\tau^{-1}\xi'$ there will be no incoming nor outgoing modes; for small values there are n incoming and n outgoing modes. The situation where the vertical line in Figure 3.3 is tangent to the slowness surface corresponds to rays tangent to the interface. Such rays are associated

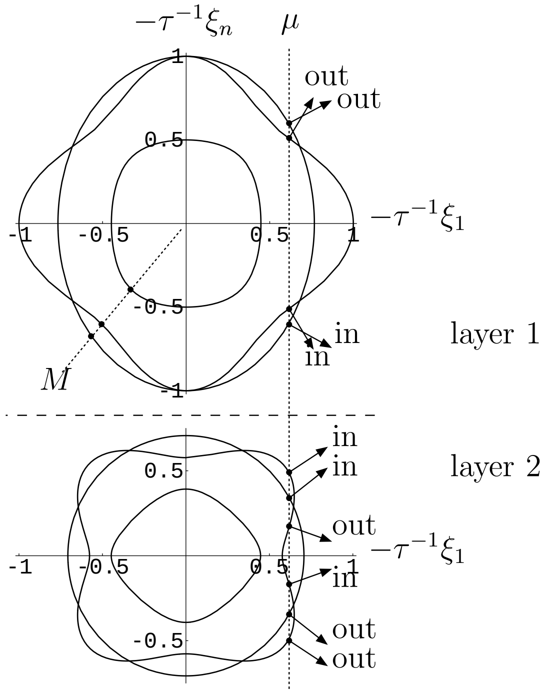


Figure 3.3: 2-dimensional section of an $n = 3$-dimensional slowness surfaces at some point of the interface, for the medium on both sides of the interface. The slownesses of the modes that interact (i.e. reflect and transmit into each other) are the intersection points with a line that is parallel to the normal of the interface. The group velocity, which is normal to the slowness surface, determines whether the mode is incoming or outgoing.

with head-waves and are not treated in our analysis. Equation (3.9) implies that the incoming and the outgoing modes correspond to the real solutions ξ_n of

$$\det P_{il}(x', 0, \xi', \xi_n, \tau) = 0.$$

This equation has $2n$ real or complex conjugated roots. The complex roots correspond to 'evanescent' wave constituents. To number the roots we use an index μ.

In the following theorem we show that if none of the rays involved is tangent, there exists a pseudodifferential operator type relation between the different modes restricted to the surface $x_n = 0$; we calculate its principal symbol in the proof. Let $x \mapsto z(x) : \mathbb{R}^n \to \mathbb{R}^n$ be a coordinate transformation such that the interface is given by $z_n = 0$. The corresponding cotangent vector is denoted by ζ, and satisfies $\zeta_i(\xi) = \sum_j \left(\frac{\partial z}{\partial x}\right)^{-1}_{ji} \xi_j$.

Assumption 2 *There are no rays tangent to the interface $z_n = 0$ microlocally at (z', t, ζ', τ).*

Theorem 3.3.1 *Suppose the roots τ of (3.9) have constant multiplicity and Assumption 2 is valid microlocally on some neighborhood in $T^*(Z' \times \mathbb{R})\backslash 0$. Let $u^{\mathrm{in}}_{N(\nu)}$ be microlocal constituents of a solution describing the 'incoming' modes, and suppose $G_{M(\mu)}$ refers to an 'outgoing' Green's function (3.19). Microlocally, the single reflected/transmitted constituent of the solution related to $u^{\mathrm{in}}_{N(\nu)}$ is given by*

$$u_{M(\mu)}(x,t) = \int_{z_n=0} G_{M(\mu)}(x, x(z), t-t_0) 2\mathrm{i} D_{t_0} \left(R_{\mu\nu}(z, D_{z'}, D_{t_0}) u^{\mathrm{in}}_{N(\nu)}(x(z), t_0) \right) \times \left| \det \frac{\partial x}{\partial z} \right| \left\| \frac{\partial z_n}{\partial x} \right\| \mathrm{d}z' \, \mathrm{d}t_0, \tag{3.28}$$

where $R_{\mu\nu}(z, D_{z'}, D_t)$ is a pseudodifferential operator of order 0.

In the proof we derive the explicit form of $R^{\mathrm{prin}}_{\mu\nu}(z, \zeta', \tau)$, see Remark 3.3.2 below. The integral $\left|\det \frac{\partial x}{\partial z}\right| \left\|\frac{\partial z_n}{\partial x}\right\| \mathrm{d}z'$ is the surface integral over the surface $z_n = 0$ with Euclidean measure in x.

Proof For the moment we assume $z(x) = x$, i.e. that we have a reflector at $x_n = 0$, and smooth coefficients on either side. We show that at the interface there is a relation of the type

$$u^{\mathrm{out}}_{M(\mu)}(x', 0, t) = R^0_{\mu\nu}(x', 0, D', D_t) u^{\mathrm{in}}_{N(\nu)}. \tag{3.29}$$

We will use the notation $c_{jk;il} = c_{ijkl}$ and also $(c_{jk})_{il} = c_{ijkl}$. The partial differential equation (3.1) reads in this notation

$$\sum_l \left(\rho \delta_{il} \frac{\partial^2}{\partial t^2} - \sum_{j,k} c_{jk;il} \frac{\partial^2}{\partial x_j \partial x_k} \right) (\rho^{-1/2} u_l) + \text{l.o.t.} \ = 0.$$

This equation can be rewritten as a first-order system in the variable x_n for the vector V_a of length $2n$ that contains both the displacement and the normal traction (normal to the surface $x_n =$ constant)

$$V_a = \begin{pmatrix} \rho^{-1/2} u_i \\ \sum_{k,l} c_{nk;il} \frac{\partial(\rho^{-1/2} u_l)}{\partial x_k} \end{pmatrix}, \qquad i = 1, \ldots, n \tag{3.30}$$

in preparation for the boundary value problem (3.26), (3.27). Here, a is an index in $\{1, \ldots, 2n\}$. The first-order system then is

$$\frac{\partial V_a}{\partial x_n} = \mathrm{i} \sum_b C_{ab}(x, D', D_t) V_b,$$

where C_{ab} is a matrix partial differential operator given to highest order by

$$C_{ab}(x, D', D_t) = -\mathrm{i} \begin{pmatrix} -\sum_{q=1}^{n-1} \sum_{j=1}^{n} (c_{nn})^{-1}_{ij} c_{nq;jl} \frac{\partial}{\partial x_q} & (c_{nn})^{-1}_{il} \\ -\sum_{p,q=1}^{n-1} b_{pq;il} \frac{\partial^2}{\partial x_p \partial x_q} + \rho \delta_{il} \frac{\partial^2}{\partial t^2} & -\sum_{p=1}^{n-1} \frac{\partial}{\partial x_p} c_{pn;ij} (c_{nn})^{-1}_{jl} \end{pmatrix}_{ab}.$$

Here, $b_{pq;il} = c_{pq;il} - \sum_{j,k=1}^{n} c_{pn;ij} (c_{nn})_{jk}^{-1} c_{nq;kl}$.

The next step is to decouple this first-order system microlocally similarly as in Section 3.2.1. This means that we want to find scalar pseudodifferential operators $C_\mu(x, D', D_t)$ and a matrix pseudodifferential operator $L_{a\mu}(x, D', D_t)$ such that

$$C_{ab}(x, D', D_t) = \sum_{\mu,\nu} L_{a\mu}(x, D', D_t) \operatorname{diag}(C_\mu(x, D', D_t))_{\mu\nu} L_{\nu b}^{-1}(x, D', D_t).$$

The principal symbols $C_\mu^{\text{prin}}(x, \xi', \tau)$ are the solutions for ξ_n of

$$\det P_{il}^{\text{prin}}(x, (\xi', \xi_n), \tau) = 0. \tag{3.31}$$

In fact it is sufficient if the transformed operator (the matrix $\operatorname{diag}(C_\mu(x, D', D_t))_{\mu\nu}$) is blockdiagonal with a block for each different real root of (3.31), a block with eigenvalues with positive imaginary part, and a block with eigenvalues with negative imaginary part. This has also been discussed by Taylor [56]. Under the assumptions of the lemma this situation can be obtained, since when varying ξ', τ, the multiplicity of a real eigenvalue only changes when the multiplicity of the corresponding root of (3.9) changes, or when two real eigenvalues become complex. The number of complex eigenvalues with positive or negative imaginary part changes only when two real eigenvalues become complex or vice versa. The latter case occurs only when there are tangent rays, and is hence excluded. The $2n \times 2n$ principal symbol $L_{a\mu}^{\text{prin}}$ (the columns appropriately normalized) is given by

$$L_{a\mu}^{\text{prin}}(x, \xi', \tau) = \begin{pmatrix} Q_{iM(\mu)}^{\text{prin}}(x, (\xi', C_\mu^{\text{prin}}(x, \xi', \tau))) \\ \sum_{k,l} c_{inkl}(-\mathrm{i}(\xi', C_\mu^{\text{prin}}(x, \xi', \tau))_k) Q_{lM(\mu)}^{\text{prin}}(x, (\xi', C_\mu^{\text{prin}}(x, \xi', \tau))) \end{pmatrix}_{a\mu}.$$

(The polarization vector $Q_{iM}(x, \xi)$ can also be defined for complex ξ). We define $V_\mu = \sum_a L(x, D', D_t)_{\mu a}^{-1} V_a$. (The index mapping $\mu \mapsto M(\mu)$ assigns the appropriate mode to the normal component of the wave vector).

If the principal symbol of $C_\mu(x, \xi', \tau)$ is real, the decoupled equation for mode μ is of hyperbolic type. It corresponds to an outgoing wave or to an incoming wave, depending on the direction of the corresponding ray. If the principal symbol of $C_\mu(x, \xi', \tau)$ is complex, the decoupled operator for mode μ is of elliptic type. Depending on the sign of the imaginary part it corresponds to a mode that grows in the n-direction, a backward parabolic equation, or one that decays, a forward parabolic equation. The growing mode has to be absent, see for instance Hörmander [27], Section 20.1.

The matrix $L_{a\mu}$ is fixed up to normalization of its columns. For the elliptic modes ($\operatorname{Im} C_\mu^{\text{prin}}(x, \xi', \tau) \neq 0$) the normalization is unimportant. For the hyperbolic modes the normalization can be such that the vector $V_\mu = \sum_a L(x, D', D_t)_{\mu a}^{-1} V_a$ agrees microlocally with the corresponding mode $u_{M,\pm}$ defined in Section 3.2. To see this assume V_μ refers to the same mode as $u_{M,\pm}$. In that case there is an invertible pseudodifferential operator $\psi(x, D, D_t)$ of order 0 such that $V_\mu = \psi u_{M,\pm}$. Now we can define $V_{\mu,\text{new}} = \psi^{-1} V_{\mu,\text{old}}$. Because ψ may depend on ξ_n, this factor cannot directly be absorbed in L. However, since $V_{\mu,\text{old}}$ satisfies a first-order hyperbolic equation the dependence on ξ_n can be eliminated and the factor ψ^{-1} can be absorbed in L.

In this proof let the in-modes be the modes for which the amplitude is known, that is the incoming hyperbolic and the growing elliptic modes. Denote by $L^{(1)}_{a\mu}, L^{(2)}_{a\mu}$ the matrix $L_{a\mu}$ on either side of the interface. We define the $2n \times 2n$ matrix L^{in} such that it contains the columns related to incoming modes out of $L^{(1)}_{a\mu}, L^{(2)}_{a\mu}$, i.e.

$$L^{\text{in}}_{a\mu} = \begin{pmatrix} L^{(1),\text{in}} & -L^{(2),\text{in}} \end{pmatrix}_{a\mu},$$

and define $L^{\text{out}}_{a\mu}$ similarly (so, here, μ is slightly different). The interface conditions (3.27) now read

$$\sum_\mu (L^{\text{out}}_{a\mu} V^{\text{out}}_\mu + L^{\text{in}}_{a\mu} V^{\text{in}}_\mu) = 0.$$

If we set $R^0_{\mu\nu} = -\sum_a (L^{\text{out}})^{-1}_{\mu a} L^{\text{in}}_{a\nu}$ (for the question whether the inverse exists, see the remark after the proof) then the part referring to the hyperbolic modes gives (3.29).

By (3.29) the u^{out}_M are determined at the interface; finding how they propagate away from the interface is a (microlocal) initial value problem similar to the problem for $G_{M,\pm}$ above, where now the x_n variable plays the role of time. The solution is again a Fourier integral operator, with canonical relation generated by the bicharacteristics. It follows that we can use $\phi_{M,\pm}(x, t - t_0, x_0, \xi_J, \tau)$ as phase function (take care that $n \notin J$). The amplitude $\mathcal{A}_{M,\pm}(x_I, x_0, \xi_J, \tau)$ satisfies the transport equation as before. However, the restriction of the Fourier integral operator to the 'initial surface' $x_n = 0$ so constructed is a pseudodifferential operator that is not necessarily the identity. Let us assume

$$u^{\text{out}}_M(x,t) = \int_{x_{0,n}=0} G_{M,\pm}(x, (x'_0, 0), t - t_0) \psi(x, D_{x'_0}, D_{t_0}) u^{\text{out}}_M(x'_0, 0, t_0) \, \mathrm{d}x'_0 \, \mathrm{d}t_0, \quad (3.32)$$

where $\psi(x, D', D_t)$ is to be found such that the restriction of this representation to $x_n = 0$ is the identity. The $\pm$ sign is chosen such that $G_{M,\pm}$ is the outgoing mode. We can use again Section 8.6 of Treves [58] to find that the principal symbol of this pseudodifferential operator should be

$$\psi(x, \xi', \tau) = \left| \frac{\partial B_M}{\partial \xi_n}(x, \xi', C^{\text{prin}}_\mu(\xi', \tau)) \right| = \left| \frac{\partial x_{M,n}}{\partial t}(x, \xi', C^{\text{prin}}_\mu(\xi', \tau), 0) \right|, \qquad (3.33)$$

i.e. the normal component of the velocity of the ray, the group velocity. We now replace $G_{M,\pm}$ by (the relevant part of) G_M, using that $G_M = \frac{1}{2}\mathrm{i} G_{M,+} B_M(x,D)^{-1} - \frac{1}{2}\mathrm{i} G_{M,-} B_M(x,D)^{-1}$. Taking this into account, and the fact that $B^{\text{prin}}_M(x,\xi) = \mp\tau$, we have now obtained (3.28) for the case that $z = x$ (no coordinate transformation).

We argue that (3.28) is also true when $z(x)$ is a general coordinate transformation. This follows from transforming the equations (3.26), (3.27) to z coordinates. To highest order the symbol of (pseudo)differential operators transforms as

$\psi^{\mathrm{transf}}(z, \zeta, \tau) = \psi(x(z), \left(\frac{\partial z}{\partial x}\right)^{\mathrm{t}} \zeta, \tau)$. Tracing the steps of the proof we find the following equivalent of (3.29)

$$u^{\mathrm{out}}_{M(\mu)}(x(z', 0), t) = R^0_{\mu\nu}(z', 0, D_{z'}, D_t) u^{\mathrm{in}}_{N(\nu)}(x(z', 0), t). \tag{3.34}$$

When the interface is at $z_n = 0$ we can obtain (3.32) in z coordinates instead of x coordinates. Transforming G_M, u_M back to x coordinates we find that for x away from the interface

$$u_M(x) = \int_{z_n=0} G_M(x, x(z), t - t_0) \left| \frac{\partial z_{M,n}}{\partial t}(z, D_{z'}, D_{t_0}) \right| u^{\mathrm{out}}_M(x(z), t_0) \left| \det \frac{\partial x}{\partial z} \right| \,\mathrm{d}z' \,\mathrm{d}t_0.$$

Here $\left| \frac{\partial z_{M,n}}{\partial t}(z, D_{z'}, D_t) \right|$ is the transformed version of (3.33). Thus expression (3.28) follows, with

$$R_{\mu\nu}(z, \zeta', \tau) = \left| \frac{\partial z_{M,n}}{\partial t}(z, \zeta', \tau) \right| \left\| \frac{\partial z_n}{\partial x} \right\|^{-1} R^0_{\mu\nu}(z, \zeta', \tau).$$

□

Remark 3.3.2 The principal symbol $R^{0,\mathrm{prin}}_{\mu\nu}(z, \zeta', \tau)$ that occurs in the proof is simply the reflection coefficient for the amplitudes. The principal symbol $R^{\mathrm{prin}}_{\mu\nu}(z, \zeta', \tau)$ is obtained by multiplying $R^{0,\mathrm{prin}}_{\mu\nu}$ with the normal component of the velocity of the ray, given (for $z(x) = x$) by (3.33). The reflection coefficients satisfy unitary relations, see Chapman [14] and Kennett [31] (the appendix to Chapter 5). These follow essentially from conservation of energy. It follows that the matrix of reflection coefficients is well defined and in particular that the inverse of $L^{\mathrm{out}}_{a\mu}$ exists. Chapman [14] also gives a direct proof of the reciprocity relations for the reflection coefficients.

Remark 3.3.3 We have shown that the reflected/transmitted wave is given by a composition of Fourier integral operators acting on the source. In the case of multiple reflections or transmissions (for instance in a medium consisting of a number of smooth domains separated by smooth interfaces) this is also the case (cf. Frazer and Sen [21]). It follows that microlocally the solution operator describing the reflected solutions is itself a Fourier integral operator, where the canonical relation is given by the generalized bicharacteristics (i.e. the reflected and transmitted bicharacteristics) and the amplitude is essentially the product of the ray amplitudes and the reflection/transmission coefficients. The integration over z' accounts for the effects associated with the interface's curvature.

3.4 The Born approximation

We discuss the modeling and inversion of seismic data in the Born approximation. The medium parameters are written as the sum of a smooth background and a singular perturbation. This is important in its own right, and it will also be a motivation for our approach to the model with smooth jumps described in the previous section.

The Born approximation has been discussed by a number of authors. In the acoustic case, allowing for multipathing (caustics), see Hansen [22] and Ten Kroode et al. [34]. For the acoustic problem with nonmaximal acquisition geometry, see Nolan and Symes [41]. For the elastic case with maximal acquisition geometry (and from a more applied point of view), see De Hoop and Brandsberg-Dahl [24]. We extend their results, and give an efficient, novel presentation. Also, we discuss in detail the different assumptions that are needed for the modeling and inversion of seismic data.

3.4.1 Modeling: Perturbation of the Green's function

In the Born approximation one assumes that the total value of the medium parameters c_{ijkl}, ρ can be written as the sum of a smooth background constituent $\rho(x), c_{ijkl}(x)$ and a singular perturbation $\delta\rho, \delta c_{ijkl}$, viz.

$$c_{ijkl} + \delta c_{ijkl}, \qquad \rho + \delta\rho.$$

This decomposition induces a perturbation of P_{il} (cf. (3.4)),

$$\delta P_{il} = \delta_{il} \frac{\delta\rho}{\rho} \frac{\partial^2}{\partial t^2} - \sum_{j,k} \frac{\partial}{\partial x_j} \frac{\delta c_{ijkl}}{\rho} \frac{\partial}{\partial x_k}.$$

We denote the causal Green's operator associated with (3.2) by G_{il} and its distribution kernel by $G_{il}(x, x_0, t - t_0)$. The first-order perturbation δG_{il} of G_{il} is derived by demanding that the first-order term in $\sum_j (P_{ij} + \delta P_{ij})(G_{jk} + \delta G_{jk})$ vanishes. This results in the representation

$$\delta G_{il}(\hat{x}, \tilde{x}, t) = -\sum_{j,k} \int_0^t \int_X G_{ij}(\hat{x}, x_0, t - t_0)\, \delta P_{jk}(x_0, D_{x_0}, D_{t_0})\, G_{kl}(x_0, \tilde{x}, t_0)\, \mathrm{d}x_0\, \mathrm{d}t_0. \tag{3.35}$$

Here, $\tilde{x}$ denotes a source location, $\hat{x}$ a receiver location, and x_0 a scattering point. Because the background model is smooth the operator δG_{il} contains only the single scattered field.

We use the decoupled equations (3.7). Omitting the pseudodifferential operators $Q_{iM}(\hat{x}, D_{\hat{x}}), Q(\tilde{x}, D_{\tilde{x}})^{-1}_{Nl}$ at the beginning and end of the product, we obtain an expression for the perturbation of the Green's function $\delta G_{MN}(\hat{x}, \tilde{x}, t)$ for the pair of modes M (scattered) and N (incident)

$$\begin{aligned} \delta G_{MN}(\hat{x}, \tilde{x}, t) = &-\sum_{i,l} \int_0^t \int_X G_M(\hat{x}, x_0, t - t_0) Q(x_0, D_{x_0})^{-1}_{Mi} \\ &\times \left(\delta_{il} \frac{\partial}{\partial t_0} \frac{\delta\rho}{\rho} \frac{\partial}{\partial t_0} - \sum_{j,k} \frac{\partial}{\partial x_{0,j}} \frac{\delta c_{ijkl}}{\rho} \frac{\partial}{\partial x_{0,k}} \right) Q(x_0, D_{x_0})_{lN} G_N(x_0, \tilde{x}, t_0)\, \mathrm{d}x_0\, \mathrm{d}t_0. \end{aligned} \tag{3.36}$$

Microlocally we can write G_M as in (3.24), with appropriate substitutions for its arguments. For G_N we use in addition the reciprocity relation $G_N(x_0, \tilde{x}, t_0) = G_N(\tilde{x}, x_0, t_0)$. The product of operators $G_M Q(x_0, D_{x_0})^{-1}_{Mi} \frac{\partial}{\partial x_{0,j}}$ is a Fourier integral operator with the same phase as G_M, and amplitude that to highest order equals the product $\mathcal{A}_M(\hat{x}_{\hat{I}}, x_0, \hat{\xi}_{\hat{J}}, \tau) Q(x_0, \hat{\xi}_0)^{-1}_{Mi} \mathrm{i}\hat{\xi}_{0,j}$, where $\hat{\xi}_0 = \xi_0(\hat{x}_{\hat{I}}, x_0, \hat{\xi}_{\hat{J}}, \tau)$. Assuming that the medium perturbation vanishes around $\hat{x}$ and $\tilde{x}$ a cutoff is introduced for t_0 near 0 and t. In the resulting expression one of the two frequency variables $\hat{\tau}, \tilde{\tau}$ can now be eliminated using the integral over t_0 (see for instance Duistermaat [17], Section 2.3). In this case the result can be obtained readily by noting that the integral over t_0 can be extended to the whole of $\mathbb{R}$ (the phase is not stationary for t_0 outside $[0, t]$), and then using that $\int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}t_0(\hat{\tau}-\tilde{\tau})} \,\mathrm{d}t_0 = 2\pi\delta(\hat{\tau} - \tilde{\tau})$. The resulting formula for δG_{MN} is, modulo lower order terms in the amplitude,

$$\begin{aligned}
\delta G_{MN}(\hat{x}, \tilde{x}, t) &= (2\pi)^{-\frac{3n+1}{4} - \frac{|\hat{J}|+|\tilde{J}|+1}{2}} \int \mathcal{B}_{MN}(\hat{x}_{\hat{I}}, \hat{\xi}_{\hat{J}}, \tilde{x}_{\tilde{I}}, \tilde{\xi}_{\tilde{J}}, x_0, \tau) \\
&\times \left(\sum_{i,j,k,l} w_{MN;ijkl}(\hat{x}_{\hat{I}}, \tilde{x}_{\tilde{I}}, x_0, \hat{\xi}_{\hat{J}}, \tilde{\xi}_{\tilde{J}}, \tau) \frac{\delta c_{ijkl}(x_0)}{\rho(x_0)} + w_{MN;0}(\hat{x}_{\hat{I}}, \tilde{x}_{\tilde{I}}, x_0, \hat{\xi}_{\hat{J}}, \tilde{\xi}_{\tilde{J}}, \tau) \frac{\delta\rho(x_0)}{\rho(x_0)} \right) \\
&\times \mathrm{e}^{\mathrm{i}\Phi_{MN}(\hat{x},\tilde{x},t,x_0,\hat{\xi}_{\hat{J}},\tilde{\xi}_{\tilde{J}},\tau)} \,\mathrm{d}x_0 \,\mathrm{d}\hat{\xi}_{\hat{J}} \,\mathrm{d}\tilde{\xi}_{\tilde{J}} \,\mathrm{d}\tau. \qquad (3.37)
\end{aligned}$$

Here (see (3.18) for the construction of ϕ_M, ϕ_N),

$$\Phi_{MN}(\hat{x}, \tilde{x}, t, x_0, \hat{\xi}_{\hat{J}}, \tilde{\xi}_{\tilde{J}}, \tau) = \phi_M(\hat{x}, x_0, t, \hat{\xi}_{\hat{J}}, \tau) + \phi_N(\tilde{x}, x_0, t, \tilde{\xi}_{\tilde{J}}, \tau) - \tau t. \qquad (3.38)$$

The amplitude factors $\mathcal{B}_{MN}$ are given by

$$\mathcal{B}_{MN}(\hat{x}_{\hat{I}}, \tilde{x}_{\tilde{I}}, x_0, \hat{\xi}_{\hat{J}}, \tilde{\xi}_{\tilde{J}}, \tau) = (2\pi)^{-\frac{n-1}{4}} \mathcal{A}_M(\hat{x}_{\hat{I}}, x_0, \hat{\xi}_{\hat{J}}, \tau) \, \mathcal{A}_N(\tilde{x}_{\tilde{I}}, x_0, \tilde{\xi}_{\tilde{J}}, \tau). \qquad (3.39)$$

We will refer to the factors $w_{MN;ijkl}, w_{MN;0}$ as the radiation patterns. They are given by

$$\begin{aligned}
w_{MN;ijkl}(\hat{x}_{\hat{I}}, \tilde{x}_{\tilde{I}}, x_0, \hat{\xi}_{\hat{J}}, \tilde{\xi}_{\tilde{J}}, \tau) &= Q_{iM}(x_0, \hat{\xi}_0) Q_{lN}(x_0, \tilde{\xi}_0) \, \hat{\xi}_{0,j} \tilde{\xi}_{0,k}, \qquad (3.40) \\
w_{MN;0}(\hat{x}_{\hat{I}}, \tilde{x}_{\tilde{I}}, x_0, \hat{\xi}_{\hat{J}}, \tilde{\xi}_{\tilde{J}}, \tau) &= -\, Q_{iM}(x_0, \hat{\xi}_0) Q_{iN}(x_0, \tilde{\xi}_0) \, \tau^2 \qquad (3.41)
\end{aligned}$$

where $\hat{\xi}_0 = \xi_0(\hat{x}_{\hat{I}}, x_0, \hat{\xi}_{\hat{J}}, \tau)$, $\tilde{\xi}_0 = \xi_0(\tilde{x}_{\tilde{I}}, x_0, \tilde{\xi}_{\tilde{J}}, \tau)$. The scattering is depicted in Figure 3.4.

We investigate the map $(\frac{\delta c_{ijkl}}{\rho}, \frac{\delta\rho}{\rho}) \mapsto \delta G_{MN}(\hat{x}, \tilde{x}, t)$ induced by (3.37). We use the notation C_{ϕ_M} to indicate the subset of the global canonical relation C_M that is associated to a phase function ϕ_M (cf. (3.15)).

Lemma 3.4.1 *Assume that if $(\hat{x}, \hat{t}, \hat{\xi}, \tau; x_0, \hat{\xi}_0) \in C_{\phi_M}$, $(\tilde{x}, \tilde{t}, \tilde{\xi}, \tau; x_0, \tilde{\xi}_0) \in C_{\phi_N}$ then $\hat{\xi}_0 + \tilde{\xi}_0 \neq 0$. Then the map $(\frac{\delta c_{ijkl}}{\rho}, \frac{\delta\rho}{\rho}) \mapsto \delta G_{MN}(\hat{x}, \tilde{x}, t)$ given by (3.37) is a Fourier integral operator $\mathcal{E}'(X) \to \mathcal{D}'(X \times X \times]0, T[)$. Its canonical relation is given by*

$$\begin{aligned}
\Lambda_{0,MN} = \{ &(\hat{x}, \tilde{x}, \hat{t} + \tilde{t}, \hat{\xi}, \tilde{\xi}, \tau; x_0, \hat{\xi}_0 + \tilde{\xi}_0) \mid \\
&(\hat{x}, \hat{t}, \hat{\xi}, \tau; x_0, \hat{\xi}_0) \in C_{\phi_M}, (\tilde{x}, \tilde{t}, \tilde{\xi}, \tau; x_0, \tilde{\xi}_0) \in C_{\phi_N} \}. \qquad (3.42)
\end{aligned}$$

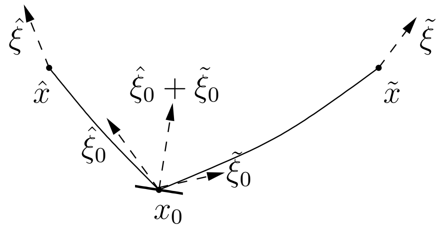


Figure 3.4: The scattering cotangent vectors.

Proof We show that $\Phi_{MN}(\hat{x}_{\hat{I}}, \tilde{x}_{\tilde{I}}, t, x_0, \hat{\xi}_{\hat{J}}, \tilde{\xi}_{\tilde{J}}, \tau)$ is a nondegenerate phase function. The derivatives with respect to the phase variables are given by

$$\begin{aligned}
\frac{\partial \Phi_{MN}}{\partial \tau} &= -\hat{t}(\hat{x}_{\hat{I}}, x_0, \hat{\xi}_{\hat{J}}, \tau) - \tilde{t}(\tilde{x}_{\tilde{I}}, x_0, \tilde{\xi}_{\tilde{J}}, \tau) + t \\
\frac{\partial \Phi_{MN}}{\partial \hat{\xi}_{\hat{J}}} &= -\hat{x}_{\hat{J}}(\hat{x}_{\hat{I}}, x_0, \hat{\xi}_{\hat{J}}, \tau) + \hat{x}_{\hat{J}} \\
\frac{\partial \Phi_{MN}}{\partial \tilde{\xi}_{\tilde{J}}} &= -\tilde{x}_{\tilde{J}}(\tilde{x}_{\tilde{I}}, x_0, \tilde{\xi}_{\tilde{J}}, \tau) + \tilde{x}_{\tilde{J}},
\end{aligned}$$

where $\hat{x}_{\hat{J}}(\hat{x}_{\hat{I}}, x_0, \hat{\xi}_{\hat{J}}, \tau), \tilde{x}_{\tilde{J}}(\tilde{x}_{\tilde{I}}, x_0, \tilde{\xi}_{\tilde{J}}, \tau)$ are as defined in (3.17), for the receiver side and the source side respectively. The derivatives of these expressions with respect to the variables $(\hat{x}_{\hat{J}}, \tilde{x}_{\tilde{J}}, t)$ are linearly independent, so Φ_{MN} is nondegenerate. From expression (3.38) it follows that the canonical relation of this operator is given by (3.42). By the assumption it contains no elements with $\hat{\xi}_0 + \tilde{\xi}_0 = 0$, so it is continuous as a map $\mathcal{E}'(X) \to \mathcal{D}'(X \times X \times]0, T[)$. $\square$

We show that the condition in Lemma 3.4.1 is violated if and only if $M = N$ and there exists a 'direct' bicharacteristic from $\tilde{x}, \tilde{\xi}$ to $\hat{x}, -\hat{\xi}$. From the symmetry of the bicharacteristic under the transformation $\xi \to -\xi, t \to -t$ it follows that indeed in this case the condition is violated. On the other hand, we have $B_M(x_0, \hat{\xi}_0) = B_N(x_0, \tilde{\xi}_0) = \pm\tau$. If $\hat{\xi}_0 = -\tilde{\xi}_0$ then we must have $M = N$, because $B_M(x_0, \hat{\xi}_0) = B_M(x_0, -\hat{\xi}_0)$ and the condition that the eigenvalues in (3.8) are different for different modes. If $M = N$ and $\hat{\xi}_0 = -\tilde{\xi}_0$ then we have the mentioned direct bicharacteristic.

3.4.2 Restriction

The data are assumed to be representable by $\delta G_{MN}(\hat{x}, \tilde{x}, t)$ for $(\hat{x}, \tilde{x}, t)$ in the acquisition manifold. To make this explicit, let $y \mapsto (\hat{x}(y), \tilde{x}(y), t(y))$ be a coordinate transformation, such that $y = (y', y'')$ and the acquisition manifold is given by $y'' = 0$. Assume that the dimension of y'' is $2 + c$, where c is the codimension of the geometry (the 2 enforces 'remote sensing'). Then the data are given by

$$\delta G_{MN}(\hat{x}(y', 0), \tilde{x}(y', 0), t(y', 0)). \tag{3.43}$$

It follows that the map $(\frac{\delta c_{ijkl}}{\rho}, \frac{\delta\rho}{\rho})$ to the data may be seen as the compose of the map of Lemma 3.4.1 with the restriction operator to $y''=0$. The restriction operator that maps a function $f(y)$ to $f(y',0)$ is a Fourier integral operator with canonical relation given by $\Lambda_r = \{(y',\eta';(y',y''),(\eta',\eta'')) \in T^*Y' \times T^*Y \,|\, y''=0\}$. The composition of the canonical relations $\Lambda_{0,MN}$ and Λ_r is well defined if the intersection of $\Lambda_r \times \Lambda_{0,MN}$ with $T^*Y'\backslash 0 \times \mathrm{diag}(T^*Y\backslash 0) \times T^*X\backslash 0$ is transversal. In this case we must have that the intersection of $\Lambda_{0,MN}$ with the manifold $y''=0$ is transversal.

Let us repeat our assumptions, and state the final result of this subsection.

Assumption 3 *There are no elements $(y',0,\eta',\eta'') \in T^*Y\backslash 0$ such that there is a direct bicharacteristic from $(\hat{x}(y',0),\hat{\xi}(y',0,\eta',\eta''))$ to $(\tilde{x}(y',0),-\tilde{\xi}(y',0,\eta',\eta''))$ with arrival time $t(y',0)$.*

Assumption 4 *The intersection of $\Lambda_{0,MN}$ with the manifold $y''=0$ is transversal. In other words,*

$$\frac{\partial y''}{\partial(x_0,\hat{\xi}_0,\tilde{\xi}_0,\hat{t},\tilde{t})} \text{ has maximal rank.} \tag{3.44}$$

In the following theorem we parameterize (3.42) by $(x_0,\hat{\xi}_0,\tilde{\xi}_0,\hat{t},\tilde{t})$ using the parameterization of C_{ϕ_M} given by (3.15). Thus we let $\tau = \mp B_M(x_0,\hat{\xi}_0)$ and

$$\begin{aligned} \hat{x} &= x_M(x_0,\hat{\xi}_0,\pm\hat{t}), & \tilde{x} &= x_N(x_0,\tilde{\xi}_0,\pm\tilde{t}), \\ \hat{\xi} &= \xi_M(x_0,\hat{\xi}_0,\pm\hat{t}), & \tilde{\xi} &= \xi_N(x_0,\tilde{\xi}_0,\pm\tilde{t}). \end{aligned}$$

We suppose that $(y'(x_0,\hat{\xi}_0,\tilde{\xi}_0,\hat{t},\tilde{t}),\eta'(x_0,\hat{\xi}_0,\tilde{\xi}_0,\hat{t},\tilde{t}))$ is obtained by transforming $(\hat{x},\tilde{x},\hat{t}+\tilde{t},\hat{\xi},\tilde{\xi},\tau)$ to (y,η) coordinates.

Theorem 3.4.2 *If Assumptions 3, 4 are satisfied then the operator $F_{MN;ijkl}$ (resp. $F_{MN;0}$) that maps the medium perturbation $\frac{\delta c_{ijkl}}{\rho}$ (resp. $\frac{\delta\rho}{\rho}$) to the data as a function of y' (3.43) is microlocally a Fourier integral operator with canonical relation given by*

$$\begin{aligned} \Lambda_{MN} = &\{(y'(x_0,\hat{\xi}_0,\tilde{\xi}_0,\hat{t},\tilde{t}),\eta'(x_0,\hat{\xi}_0,\tilde{\xi}_0,\hat{t},\tilde{t});x_0,\hat{\xi}_0+\tilde{\xi}_0) \,|\, \\ &B_M(x_0,\hat{\xi}_0) = B_N(x_0,\tilde{\xi}_0) = \pm\tau, y''(x_0,\hat{\xi}_0,\tilde{\xi}_0,\hat{t},\tilde{t}) = 0\}. \end{aligned} \tag{3.45}$$

The order equals $\frac{n-1+c}{4}$. The amplitude is given to highest order (in coordinates (y'_I,η'_J,x_0) for Λ_{MN}, where I,J is a partition of $\{1,\ldots,2n-1-c\}$) by the product $\mathcal{B}_{MN}(y'_I,\eta'_J,x_0)w_{MN;ijkl}(y'_I,\eta'_J,x_0)$ (resp. $\mathcal{B}_{MN}(y'_I,\eta'_J,x_0)w_{MN;0}(y'_I,\eta'_J,x_0)$), where

$$\begin{aligned} |\mathcal{B}_{MN}(y'_I,\eta'_J,x_0)| = &\tfrac{1}{4}\tau^{-2}(2\pi)^{-\frac{n+1+c}{4}} \\ &\times \left|\det\frac{\partial(\hat{x},\tilde{x},t)}{\partial y}\right|^{-\frac{1}{2}} \left|\det\frac{\partial(x_0,\hat{\xi}_0,\tilde{\xi}_0,\hat{t},\tilde{t})}{\partial(x_0,y'_I,y'',\eta'_J,\Delta\tau)}\right|^{\frac{1}{2}}_{\Delta\tau=0,y''=0} \end{aligned} \tag{3.46}$$

Here we define $\Delta\tau = \hat{\tau}-\tilde{\tau}$, so that the first constraint in (3.45) reads $\Delta\tau=0$. The map $(x_0,\hat{\xi}_0,\tilde{\xi}_0,\hat{t},\tilde{t}) \mapsto (x_0,y'_I,y'',\eta'_J,\Delta\tau)$ is bijective.

Proof The first statement has been argued above. The order of the operator is given by

$$\chi + \frac{K}{2} - \frac{\dim X + \dim Y'}{4},$$

where χ is the degree of homogeneity of the amplitude and K is the number of phase variables. The factors $\{w_{MN;ijkl}, w_{MN;0}\}$ are homogeneous of order 2 in the ξ and τ variables; the degree of homogeneity of the factor $\mathcal{B}_{MN}$ follows from (3.22). We find

$$\begin{aligned}\text{order } F_{MN;ijkl} &= 2 + (-2 - \frac{|\hat{J}| + |\tilde{J}| + 2}{2} + n) + \frac{|\hat{J}| + |\tilde{J}| + 1}{2} - \frac{3n - 1 - c}{4} \\ &= \frac{n - 1 + c}{4}.\end{aligned}$$

We calculate now the amplitude of the Fourier integral operator in Lemma 3.4.1. The factor $w_{MN;ijkl}$ is simply multiplicative. Suppose we choose coordinates on $\Lambda_{0,MN}$ to be $(\hat{x}_{\hat{I}}, \hat{\xi}_{\hat{J}}, \tilde{x}_{\tilde{I}}, \tilde{\xi}_{\tilde{J}}, \hat{\tau}, \tilde{\tau}, x_0)$, with ultimately $\hat{\tau} = \tilde{\tau}$. Define $\tau = \frac{\hat{\tau}+\tilde{\tau}}{2}$, $\Delta\tau = \hat{\tau} - \tilde{\tau}$. Using (3.25) and (3.39) we find that the amplitude $\mathcal{B}_{MN}(x_0, \hat{x}_{\hat{I}}, \hat{\xi}_{\hat{J}}, \tilde{x}_{\tilde{I}}, \tilde{\xi}_{\tilde{J}}, \tau)$ is given by

$$|\mathcal{B}_{MN}(\hat{x}_{\hat{I}}, \hat{\xi}_{\hat{J}}, \tilde{x}_{\tilde{I}}, \tilde{\xi}_{\tilde{J}}, \tau, x_0)| = \tfrac{1}{4}\tau^{-2}(2\pi)^{-\frac{n-1}{4}} \left|\det \frac{\partial(x_0, \hat{\xi}_0, \tilde{\xi}_0, \hat{t}, \tilde{t})}{\partial(\hat{x}_{\hat{I}}, \hat{\xi}_{\hat{J}}, \tilde{x}_{\tilde{I}}, \tilde{\xi}_{\tilde{J}}, \tau, x_0, \Delta\tau)}\right|^{\frac{1}{2}}$$

The transformation from $(\hat{x}, \tilde{x}, t)$ to y coordinates in Fourier integral (3.43), induces an additional factor $|\det \frac{\partial(\hat{x},\tilde{x},t)}{\partial y}|^{-\frac{1}{2}}$ (note that for the Fourier integral operators it would be more natural to transform as a half-density). The amplitude transforms as a half-density on the canonical relation, and we obtain the factor

$$\left|\det \frac{\partial(y'_I, y'', \eta'_J)}{\partial(\hat{x}_{\hat{I}}, \hat{\xi}_{\hat{J}}, \tilde{x}_{\tilde{I}}, \tilde{\xi}_{\tilde{J}}, \tau)}\right|^{\frac{1}{2}}.$$

The additional factor $(2\pi)^{-\frac{2+c}{4}}$ arises from the normalization. We find (3.46). $\square$

Natural coordinates for the canonical relation are given by $(x_0, \hat{\xi}_0, \tilde{\xi}_0, \hat{t}, \tilde{t})$ such that $B_M(x_0, \hat{\xi}_0) - B_N(x_0, \tilde{\xi}_0) = 0$, $y''(x_0, \hat{\xi}_0, \tilde{\xi}_0, \hat{t}, \tilde{t}) = 0$. There is a natural density directly associated with this set, the quotient density. The Jacobian in (3.46) reveals that the amplitude factor $|\mathcal{B}_{MN}(y'_I, \eta'_J, x_0)|$ is in fact given by the associated half-density times $\frac{1}{4}\tau^{-2}(2\pi)^{-\frac{n+1+c}{4}} \left|\frac{\partial(\hat{x},\tilde{x},t)}{\partial y}\right|^{-\frac{1}{2}}$.

If $c = 0$ and there are no rays tangent to the acquisition manifold, i.e.

$$\operatorname{rank} \frac{\partial y''}{\partial(\hat{t}, \tilde{t})} = 2, \tag{3.47}$$

then a convenient way to parameterize the canonical relation is found using the phase directions $\hat{\alpha} = \frac{\hat{\xi}_0}{\|\hat{\xi}_0\|}, \tilde{\alpha} = \frac{\tilde{\xi}_0}{\|\tilde{\xi}_0\|} \in S^{n-1}$ and the frequency τ.

3.4.3 Inversion

Let us now consider the reconstruction of $\left(\frac{\delta c_{ijkl}}{\rho}, \frac{\delta\rho}{\rho}\right)$ from the data. We simplify the notation, and collect the medium perturbations into

$$g_\alpha = \left(\frac{\delta c_{ijkl}}{\rho}, \frac{\delta\rho}{\rho}\right).$$

The forward operator $(F_{MN;ijkl}, F_{MN;0})$ in the Born approximation is represented by$F_{MN;\alpha}$.

Let us consider data from a single pair of modes (M, N) (the general case is discussed at the end of this section). The standard procedure to deal with the fact that this inverse problem is overdetermined is to use the method of least squares. Define the normal operator $N_{MN;\alpha\beta}$ as the product of $F_{MN;\alpha}$ and its adjoint $F^*_{MN;\alpha}$,

$$N_{MN;\alpha\beta} = F^*_{MN;\alpha} F_{MN;\beta}. \tag{3.48}$$

If $N_{MN;\alpha\beta}$ is invertible (as a matrix-valued operator with indices $\alpha\beta$), then

$$F^{-1}_{MN;\alpha} = \sum_\beta (N_{MN})^{-1}_{\alpha\beta} F^*_{MN;\beta} \tag{3.49}$$

is a left inverse of $F_{MN;\alpha}$ that is optimal in the sense of least squares[1].

The properties of the compose (3.48) depend on those of Λ_{MN}. Let $\pi_{Y'}, \pi_X$ be the projection mappings of Λ_{MN} to $T^*Y'\backslash 0$, $T^*X\backslash 0$ respectively. We will show that under the following assumption, $N_{MN;\alpha\beta}$ is a pseudodifferential operator, so that the problem of inverting $N_{MN;\alpha\beta}$ reduces to a finite-dimensional problem for each $(x, \xi) \in \pi_X(\Lambda_{MN})$.

Assumption 5 *The projection $\pi_{Y'}$ of Λ_{MN} on $T^*Y'\backslash 0$, is an embedding, i.e. it is*
i) immersive
ii) injective
iii) proper .

This assumption implies that the image of $\pi_{Y'}$ is a submanifold, $\mathcal{L}$ say, of $T^*Y'\backslash 0$. Let us discuss these requirements, starting with the first. Using that Λ_{MN} is a canonical relation we have

Lemma 3.4.3 *The projection $\pi_{Y'}$ of Λ_{MN} on $T^*Y'\backslash 0$ is an immersion if and only if the projection π_X of Λ_{MN} on $T^*X\backslash 0$ is a submersion. In this case the image of $\pi_{Y'}$ is locally a coisotropic submanifold of $T^*Y'\backslash 0$.*

[1]Equation (3.48) is for the case where one minimizes the difference with the data δG_{MN} in L^2 norm $\|\delta G_{MN} - F_{MN;\alpha} g_\alpha\|$. It can easily be adapted to the case where one minimizes a Sobolev norm of different order, or a weighted L^2 norm. This would introduce extra factors in the amplitude.

Proof This is a property of Lagrangian manifolds. It follows from Lemma 25.3.6 in Hörmander [27]. We give an independent proof.

The symplectic forms $\sigma_X, \sigma_{Y'}$ on $T^*X\backslash 0, T^*Y'\backslash 0$ can be viewed as 2-forms on Λ_{MN}. Because Λ_{MN} is a canonical relation, $\sigma_{Y'} = \sigma_X$ on Λ_{MN}, and in particular $\operatorname{rank}\sigma_{Y'} = \operatorname{rank}\sigma_X$. Now consider π_X. Clearly $\operatorname{rank}\sigma_X = 2n$ if and only if π_X is submersive.

Consider $\pi_{Y'}$. If this projection is immersive then the image has dimension $n+m$, assuming $\dim T^*Y'\backslash 0 = 2m$ (in this proof $m = \dim Y' = 2n - 1 - c$). Then $\operatorname{rank}\sigma_{Y'}$ is at least $2n$, so it must be equal to $2n$. On the other hand, if $\operatorname{rank}\sigma_{Y'} = 2n$, then the tangent space of Λ_{MN} at that point is given by the span of a set vectors of the form

$$\{(v_1, w_1), \ldots, (v_{2n}, w_{2n}), (0, w_{2n+1}), \ldots, (0, w_{n+m})\}.$$

The $w_i, i \in \{1, \ldots, 2n\}$ must be linearly independent because $\operatorname{rank}\sigma_{Y'} = 2n$. For $w_i, w_j, i \leq 2n, j > 2n$ we have $\sigma_{Y'}(w_i, w_j) = 0$, so the w_j are linearly independent from the w_i. The $w_i, i > 2n$ must be linearly independent, because $(0, w_i)$ are basis vectors for the tangent space to Λ_{MN}. So if $\operatorname{rank}\sigma_{Y'} = 2n$ then $\pi_{Y'}$ is an immersion. Because $\operatorname{rank}\sigma_{Y'} = 2n$ in that case, the image is locally a coisotropic submanifold. □

As a consequence, if part i) of Assumption 5 is satisfied then we can use $(x, \xi) \in T^*X\backslash 0$ as (local) coordinates on Λ_{MN}. In addition, we need to parameterize the subsets of the canonical relation given by $(x, \xi) =$ constant; we denote such parameters by e. The new parameterization of Λ_{MN} is (identifying x_0 with x)

$$\Lambda_{MN} = \{(y'(x, \xi, e), \eta'(x, \xi, e); (x, \xi))\}. \tag{3.50}$$

The results do not depend on the precise definition of e. As noted before, if the variables $(\hat{t}, \tilde{t})$ can be solved from the second constraint in (3.45) (cf. equation (3.47)), then Λ_{MN} can be parameterized using $(x, \hat{\alpha}, \tilde{\alpha}, \tau)$, where $(\hat{\alpha}, \tilde{\alpha})$ are phase directions. In that case (x, ξ, e) should be related by a coordinate transformation to $(x, \hat{\alpha}, \tilde{\alpha}, \tau)$. In acoustic media (where $\|\hat{\xi}_0\| = \|\tilde{\xi}_0\|$) a suitable choice is the pair scattering angle/azimuth, given by

$$\left(\arccos(\hat{\alpha} \cdot \tilde{\alpha}), \frac{-\hat{\alpha} + \tilde{\alpha}}{2\sin(\arccos(\hat{\alpha} \cdot \tilde{\alpha})/2)}\right) \in]0, \pi[\, \times S^{n-2},$$

cf. Burridge and Beylkin [13]. The azimuth, the second component, defines together with ξ the plane spanned by $(\hat{\alpha}, \tilde{\alpha})$. It is not difficult to show that in elastic media the scattering angle (the first component) can be used as coordinate when the slowness sheets are convex, but not always when one of the slowness sheets fails to be convex.

Remark 3.4.4 We show that the first part of Assumption 5 implies that $\frac{\partial B_M}{\partial \xi}(x, \hat{\xi}_0) + \frac{\partial B_N}{\partial \xi}(x, \tilde{\xi}_0) \neq 0$, in other words the group velocities at the scattering point do not add up to 0. We have seen in Theorem 3.4.2 that Λ_{MN} may be parameterized by $(x, \hat{\xi}_0, \tilde{\xi}_0, \hat{t}, \tilde{t})$, where $(\hat{\xi}_0, \tilde{\xi}_0)$ are such that

$$B_M(x_0, \hat{\xi}_0) = B_N(x_0, \tilde{\xi}_0) = \pm\tau$$

(and we have the additional constraint $y''(x_0, \hat{\xi}_0, \tilde{\xi}_0, \hat{t}, \tilde{t}) = 0$). The projection π_X is given by $(x, \hat{\xi}_0 + \tilde{\xi}_0)$. Consider tangent vectors to Λ_{MN} given by vectors $v_{\hat{\xi}_0}, v_{\tilde{\xi}_0}$. They must satisfy

$$v_{\hat{\xi}_0} \cdot \frac{\partial B_M}{\partial \xi}(x, \hat{\xi}_0) = v_{\tilde{\xi}_0} \cdot \frac{\partial B_N}{\partial \xi}(x, \tilde{\xi}_0) = \pm v_\tau. \tag{3.51}$$

So if $\frac{\partial B_M}{\partial \xi}(x, \hat{\xi}_0) = -\frac{\partial B_N}{\partial \xi}(x, \tilde{\xi}_0)$, then (3.51) implies that $(v_{\hat{\xi}_0} + v_{\tilde{\xi}_0}) \cdot \frac{\partial B_M}{\partial \xi}(x, \hat{\xi}_0) = 0$, so that the projection of Λ_{MN} on $T^*X\backslash 0$ is not submersive. If $c = 0$, and rank $\frac{\partial y''}{\partial(\hat{t}, \tilde{t})} = 2$ (no tangent rays), then the constraint $y'' = 0$ may be used to solve for the parameters $\hat{t}, \tilde{t}$ and (3.51) is the only condition on $(\hat{\xi}_0, \tilde{\xi}_0)$. In that case $\frac{\partial B_M}{\partial \xi}(x, \hat{\xi}_0) \neq -\frac{\partial B_N}{\partial \xi}(x, \tilde{\xi}_0)$ implies that the projection is submersive. In other cases the set of $(\hat{\xi}_0, \tilde{\xi}_0)$ is in general a smaller subset of $T_x^*X\backslash 0 \times T_x^*X\backslash 0$.

Let us now discuss the second and third parts of Assumption 5. The second part is a well known condition, see Hansen [22] and Ten Kroode et al. [34]. Essentially the condition is that there are no two different singularities in g_α mapped to the same position in $T^*Y'\backslash 0$. For an analysis of the case where this condition is violated, see Stolk [52].

The definition of proper is that the preimage of a compact set is a compact set. So assume we have a compact subset of $T^*Y'\backslash 0$. The elements of Λ_{MN} correspond to those 'points' where the source and receiver rays intersect. The set of these points can be written as a set on which some continuous function vanishes. Therefore this set is closed. It is also bounded, and hence it is compact. So the third part of the assumption is automatically satisfied.

When constructing the compose (3.48) there is a subtlety that we have to take into account, namely that the linearized forward operator is only *microlocally* a Fourier integral operator. To make it globally a Fourier integral operator, we apply a pseudodifferential cutoff $\psi(y', D_{y'})$ with compact support. Due to the third part of Assumption 5, the forward operator is then a finite sum of local Fourier integral operators.

Theorem 3.4.5 *Let $\psi(y', D_{y'})$ be a pseudodifferential cutoff with conically compact support in $T^*Y'\backslash 0$, such that for the set*

$$\{(y', \eta'; x_0, \xi_0) \in \Lambda_{MN} \mid (y', \eta') \in \operatorname{supp} \psi\} \tag{3.52}$$

Assumptions 3, 4, 5 are satisfied. Then

$$F^*_{MN;\beta}\psi(y', D_{y'})^*\psi(y', D_{y'})F_{MN;\alpha} \tag{3.53}$$

is a pseudodifferential operator of order $n-1$. Its principal symbol is given by

$$N_{MN;\beta\alpha}(x,\xi) = \frac{1}{16}(2\pi)^{-n}\int |\psi(y'(x,\xi,e),\eta'(x,\xi,e))|^2\tau^{-4}\overline{w_{MN;\beta}(x,\xi,e)}w_{MN;\alpha}(x,\xi,e)$$
$$\times\left|\det\frac{\partial(\hat{x},\tilde{x},t)}{\partial y}\right|^{-1}\left|\det\frac{\partial(x,\hat{\xi}_0,\tilde{\xi}_0,\hat{t},\tilde{t})}{\partial(x,\xi,e,y'',\Delta\tau)}\right|_{\Delta\tau=0,y''=0}\mathrm{d}e, \tag{3.54}$$

where $\tau=\tau(x,\xi,e)$.

Proof We use the clean intersection calculus for Fourier integral operators (see e.g. Treves [58]) to show that (3.53) is a Fourier integral operator. The canonical relation of $F^*_{MN;\alpha}$ is given by

$$\Lambda^*_{MN} = \{(x,\xi;y',\eta')\,|\,(y',\eta';x,\xi)\in\Lambda_{MN}\}.$$

Let $L=\Lambda^*_{MN}\times\Lambda_{MN}$ and $M=T^*X\backslash 0\times\operatorname{diag}(T^*Y'\backslash 0)\times T^*X\backslash 0$. We have to show that the intersection of $L\cap M$ is clean, i.e.

$$L\cap M \text{ is a manifold,} \tag{3.55}$$
$$TL\cap TM = T(L\cap M). \tag{3.56}$$

It follows from Assumption 5 ii) that $L\cap M$ must be given by

$$L\cap M = \{(x,\xi,y',\eta',y',\eta',x,\xi)\,|\,(y',\eta';x,\xi)\in\Lambda_{MN}\}. \tag{3.57}$$

Because Λ_{MN} is a manifold this set satisfies (3.55). The property (3.56) follows from the assumption that the map $\pi_{Y'}$ is immersive. The excess is given by

$$\begin{aligned} E =\ & \dim(L\cap M) - (\dim L + \dim M - \dim T^*X\backslash 0\times T^*Y'\backslash 0\times T^*Y'\backslash 0\times T^*X\backslash 0)\\ =\ & n-1-c. \end{aligned} \tag{3.58}$$

Taking into account the pseudodifferential cutoff $\psi(y',D_{y'})$, it follows that (3.53) is a Fourier integral operator. The canonical relation $\Lambda^*_{MN}\circ\Lambda_{MN}$ of $F^*_{MN;\beta}\psi^*\psi F_{MN;\alpha}$ is contained in the diagonal of $T^*X\backslash 0\times T^*X\backslash 0$, so it is a pseudodifferential operator. The order is given by $2\operatorname{order}F_{MN;\alpha}+\frac{E}{2}=n-1$ (note that c drops out).

We write $\psi(y',D_{y'})^*\psi(y',D_{y'}) = \sum_i\chi^{(i)}(y',D_{y'})$, where the symbols $\chi^{(i)}(y',\eta')$ have small enough support, so that the distribution kernel of $\chi^{(i)}(y',D_{y'})F_{MN;\alpha}$ can be written as the oscillatory integral

$$\chi^{(i)}(y',D_{y'})\mathcal{F}_{MN;\alpha}(y',x) = (2\pi)^{-\frac{3n-1-c}{4}-\frac{|J|}{2}}\int\chi^{(i)}(y'_I,\eta'_J,x)$$
$$\times\mathcal{B}_{MN}(y'_I,\eta'_J,x)w_{MN;\alpha}(y'_I,\eta'_J,x)\mathrm{e}^{\mathrm{i}(S^{(i)}_{MN}(y'_I,x,\eta'_J)+\langle\eta'_J,y'_J\rangle)}\,\mathrm{d}\eta'_J, \tag{3.59}$$

where $\psi^{(i)}(y'_I,\eta'_J,x)=\psi^{(i)}(y'_I,y'_J(y'_I,\eta'_J,x),\eta'_I(y'_I,\eta'_J,x),\eta'_J)$, and we used that we can write $\Phi^{(i)}_{MN}(y',x,\eta'_J)=S^{(i)}_{MN}(y'_I,x,\eta'_J)+\langle\eta'_J,y'_J\rangle$, (cf. (3.18),(3.38)). We do not indicate the dependence of J on i explicitly. The distribution kernel of the normal operator is then given by a sum of terms

$$\begin{aligned}
&\int (\overline{\psi(y', D_{y'})\mathcal{F}_{MN;\beta}(y',x)})(\psi(y', D_{y'})\mathcal{F}_{MN;\alpha}(y',x_0))\,\mathrm{d}y' \\
&= (2\pi)^{-\frac{3n-1-c}{2}-|J|} \sum_i \int \chi^{(i)}(y'_I, \eta'_{0,J}, x_0) \\
&\quad\times \overline{\mathcal{B}_{MN}(y'_I, \eta'_J, x)}\mathcal{B}_{MN}(y'_I, \eta'_{0,J}, x_0)\overline{w_{MN;\beta}(y'_I, \eta'_J, x)}w_{MN;\alpha}(y'_I, \eta'_{0,J}, x_0) \\
&\quad\times \mathrm{e}^{\mathrm{i}(S^{(i)}_{MN}(y'_I, x_0, \eta'_{0,J}) - S^{(i)}_{MN}(y'_I, x, \eta'_J) + \langle \eta'_{0,J}, y'_J\rangle - \langle \eta'_J, y'_J\rangle)}\,\mathrm{d}\eta'_{0,J}\,\mathrm{d}\eta'_J\,\mathrm{d}y'.
\end{aligned}$$

We now apply the method of stationary phase, and integrate out the variables $y'_J, \eta'_{0,J}$. For the remaining variables we use that

$$S^{(i)}_{MN}(y'_I, x_0, \eta'_J) - S^{(i)}_{MN}(y'_I, x, \eta'_J) = \langle x - x_0, \xi(y'_I, \eta'_J, x_0)\rangle + O(|x - x_0|^2).$$

Thus we find (to highest order)

$$\begin{aligned}
&(2\pi)^{-\frac{3n-1-c}{2}} \sum_i \int \chi^{(i)}(y'_I, \eta'_J, x)^2 |\mathcal{B}_{MN}(y'_I, \eta'_J, x)|^2 \overline{w_{MN;\beta}(y'_I, \eta'_J, x)} w_{MN;\alpha}(y'_I, \eta'_J, x) \\
&\quad\times \mathrm{e}^{\mathrm{i}\langle x - x_0, \xi(y'_I, \eta'_J, x_0)\rangle}\,\mathrm{d}\eta'_J\,\mathrm{d}y'_I.
\end{aligned}$$

We now change of variables $(x, y'_I, \eta'_J) \to (x, \xi, e)$, and use (3.46). We sum over i, and arrive at

$$\begin{aligned}
\mathcal{N}_{MN;\beta\alpha}(x, x_0) &= \frac{(2\pi)^{-2n}}{16} \int |\psi(y'(x,\xi,e), \eta'(x,\xi,e))|^2 \tau^{-4} \overline{w_{MN;\beta}(x,\xi,e)} w_{MN;\alpha}(x,\xi,e) \\
&\quad\times \left|\det \frac{\partial(\hat{x}, \tilde{x}, t)}{\partial y}\right|^{-1} \left|\det \frac{\partial(x, \hat{\xi}_0, \tilde{\xi}_0, \hat{t}, \tilde{t})}{\partial(x, \xi, e, y'', \Delta\tau)}\right|_{\Delta\tau=0, y''=0} \mathrm{e}^{\mathrm{i}\langle x - x_0, \xi\rangle}\,\mathrm{d}\xi\,\mathrm{d}e. \qquad (3.60)
\end{aligned}$$

It follows that the principal symbol of $N_{MN;\beta\alpha}$ is given by (3.54). $\square$

So far we focused on inversion of data from one pair of modes (M, N). Often data will be available for some subset S of all possible pairs of modes. Define the normal operator for this case as

$$N_{\alpha\beta} = \sum_{(M,N)\in S} F^*_{MN;\alpha} F_{MN;\beta} = \sum_{(M,N)\in S} N_{MN;\alpha\beta}.$$

If all the $N_{MN;\alpha\beta}$ are pseudodifferential operators then $N_{\alpha\beta}$ is also a pseudodifferential operator. A left inverse is now given by

$$N^{-1}_{\alpha\beta} F^*_\beta,$$

where F^*_β is the vector of Fourier integral operators containing the $F^*_{MN;\beta}$, $(M, N) \in S$.

3.5 Symplectic geometry of the data

In the previous section we saw that the wavefront set of the modeled data cannot be arbitrary. This is due to the redundancy in the data: in the Born approximation the singular part of the medium parameters is a function of n variables, while the data is a function of $2n-1-c$ variables. This redundancy is employed in the parameter reconstruction, and is important in the reconstruction of the background medium (or the medium above the interface in the case of a smooth jump) as well. This will be explained below.

Consider again the canonical relation Λ_{MN}. Suppose Assumption 5 is satisfied. Denote in this section by Ω the map

$$\Omega : (x, \xi, e) \mapsto (y'(x, \xi, e), \eta'(x, \xi, e)) : T^*X\backslash 0 \times E \to T^*Y'\backslash 0$$

introduced above (3.50). This map conserves the symplectic form of $T^*X\backslash 0$. That is, if $w_{x_i} = \frac{\partial(y',\eta')}{\partial x_i}$ and similarly for w_{ξ_i}, w_{e_i}, we have

$$\begin{aligned}
&\sigma_{Y'}(w_{x_i}, w_{x_j}) = \sigma_{Y'}(w_{\xi_i}, w_{\xi_j}) = 0,\\
&\sigma_{Y'}(w_{\xi_i}, w_{x_j}) = \delta_{ij},\\
&\sigma_{Y'}(w_{e_i}, w_{x_j}) = \sigma_{Y'}(w_{e_i}, w_{\xi_j}) = \sigma_{Y'}(w_{e_i}, w_{e_j}) = 0.
\end{aligned} \tag{3.61}$$

The (x, ξ, e) are 'symplectic coordinates' on the projection of Λ_{MN} on $T^*Y'\backslash 0$, which is a subset $\mathcal{L}$ of $T^*Y'\backslash 0$.

The image $\mathcal{L}$ of the map Ω is coisotropic, as noted in Lemma 3.4.3. The sets $(x, \xi) =$ constant are the isotropic fibers of the fibration of Hörmander [26], Theorem 21.2.6, see also Theorem 21.2.4. Duistermaat [17] calls them characteristic strips (see Theorem 3.6.2). We have sketched the situation in Figure 3.5. The wavefront set of the data is contained in $\mathcal{L}$ and is a union of fibers.

Using the following result we can extend the coordinates (x, ξ, e) to symplectic coordinates on an open neighborhood of $\mathcal{L}$.

Lemma 3.5.1 *Let $\mathcal{L}$ be an embedded coisotropic submanifold of $T^*Y'\backslash 0$, with coordinates (x, ξ, e) such that (3.61) holds. Denote $\mathcal{L} \ni (y', \eta') = \Omega(x, \xi, e)$. We can find a homogeneous canonical map G from an open part of $T^*(X \times E)\backslash 0$ to an open neighborhood of $\mathcal{L}$ in $T^*Y'\backslash 0$, such that $G(x, e, \xi, \epsilon = 0) = \Omega(x, \xi, e)$.*

Proof The e_i can be viewed as (coordinate) functions on $\mathcal{L}$. We will first extend them to functions on the whole $T^*Y'\backslash 0$ such that the Poisson brackets $\{e_i, e_j\}$ satisfy

$$\{e_i, e_j\} = 0, \qquad 1 \le i, j \le m - n, \tag{3.62}$$

where $m = \dim Y' = 2n-c-1$. This can be done successively for $e_1, \dots, e_{m-n}$ by the method that we describe now, see Treves [58], Chapter 7, the proof of Theorem 3.3, or

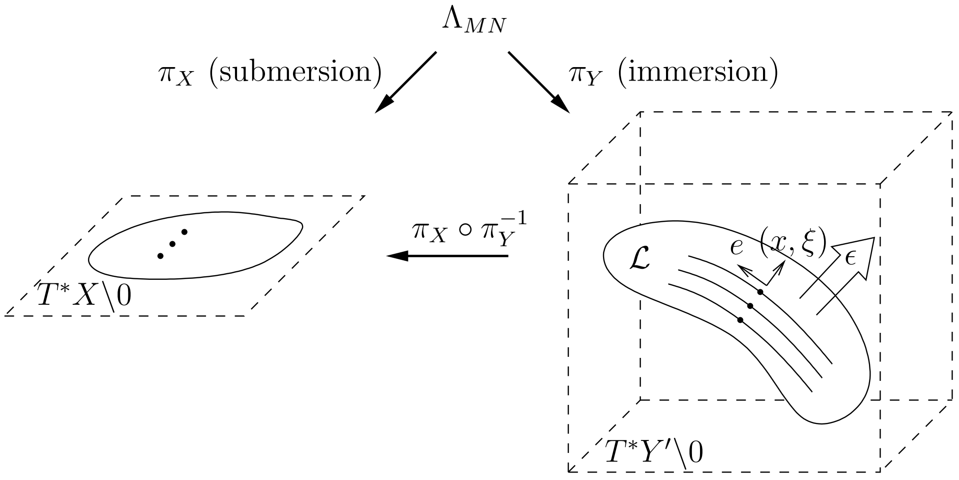


Figure 3.5: Visualization of the symplectic structure of Λ_{MN} (cone structure omitted).

Duistermaat [17], the proof of Theorem 3.5.6. Suppose we have extended $e_1, \dots, e_l$, we extend e_{l+1}. In order to satisfy (3.62) e_{l+1} has to be a solution u of

$$H_{e_i} u = 0, \qquad 1 \le i \le l,$$

where H_{e_i} is the Hamilton field associated with the function e_i, with initial condition on some manifold transversal to the H_{e_i}. For any $(y', \eta') \in \mathcal{L}$ the covectors $\mathrm{d}e_i$, $1 \le i \le l$ restricted to $T_{(y',\eta')}\mathcal{L}$ are linearly independent, so the H_{e_i} are transversal to $\mathcal{L}$ and they are linearly independent modulo $\mathcal{L}$. So we can give the initial condition $u|_{\mathcal{L}} = e_{l+1}$ and even prescribe u on a larger manifold, which leads to nonuniqueness of the extensions e_i.

We now have $m - n$ commuting vectorfields H_{e_i} that are transversal to $\mathcal{L}$ and linearly independent on some open neighborhood of $\mathcal{L}$. The Hamilton systems with parameters ϵ_i read

$$\frac{\partial y'_j}{\partial \epsilon_i} = \frac{\partial e_i}{\partial \eta'_j}(y', \eta'), \qquad \frac{\partial \eta'_j}{\partial \epsilon_i} = -\frac{\partial e_i}{\partial y'_j}(y', \eta'), \qquad 1 \le i, j \le m - n.$$

Let $G(x, e, \xi, \epsilon)$ be the solution for (y', η') of the Hamilton systems combined with initial value $(y', \eta') = \Omega(x, \xi, e)$ with 'flowout parameters' ϵ. This gives a diffeomorphic map from a neigborhood of the set $\epsilon = 0$ in $T^*(X \times E)\backslash 0$ to a neighborhood of $\mathcal{L}$ in $T^*Y'\backslash 0$. One can check from the Hamilton systems that this map is homogeneous.

It remains to check the commutation relations. The relations (3.61) are valid for any ϵ, because the Hamilton flow conserves the symplectic form on $T^*Y'\backslash 0$. The commutation relations for $\frac{\partial (y',\eta')}{\partial \epsilon_i}$ follow, using that $\frac{\partial (y',\eta')}{\partial \epsilon_i} = H_{e_i}$. $\square$

Let M_{MN} be the canonical relation associated to the map G we just constructed, i.e. $M_{MN} = \{(G(x, e, \xi, \epsilon); x, e, \xi, \epsilon)\}$. We now construct a Maslov-type phase function for M_{MN} that is directly related to a phase function for Λ_{MN}. Suppose (y'_I, η'_J, x) are suitable coordinates for Λ_{MN} ($\epsilon = 0$). For ϵ small, the constant-ϵ subset of M_{MN} can be coordinatized by the same set of coordinates, thus we can use coordinates $(y'_I, \eta'_J, x, \epsilon)$ on M_{MN}. Now there is (see Theorem 4.21 in Maslov and Fedoriuk [38]) a function $S_{MN}(y'_I, x, \eta'_J, \epsilon)$ such that M_{MN} is given by

$$\begin{aligned} y'_J &= -\frac{\partial S_{MN}}{\partial \eta'_J}, & \eta'_I &= \frac{\partial S_{MN}}{\partial y'_I}, \\ \xi &= -\frac{\partial S_{MN}}{\partial x}, & e &= \frac{\partial S_{MN}}{\partial \epsilon}. \end{aligned}$$

Thus a phase function for M_{MN} is given by

$$\Psi_{MN}(y', x, e, \eta'_J, \epsilon) = S_{MN}(y'_I, x, \eta'_J, \epsilon) + \langle \eta'_J, y'_J \rangle - \langle \epsilon, e \rangle. \tag{3.63}$$

A Maslov-type phase function for Λ_{MN} then follows as

$$\Phi_{MN}(y', x, \eta'_J) = \Psi_{MN}(y', x, \tfrac{\partial S_{MN}}{\partial \epsilon}|_{\epsilon=0}, \eta'_J, 0) = S_{MN}(y'_I, \eta'_J, x, 0) + \langle \eta'_J, y'_J \rangle.$$

In the absence of caustics there is natural choice for the symplectic coordinates given by the map G in Lemma 3.5.1, using that the time variables plays a special role. We explain this for codimension $c = 0$, where $y' = (\hat{x}, \tilde{x}, t) \in \partial X \times \partial X \times]0, T[$, and assuming that equation (3.47) is satisfied. Under a no caustics assumption $(\hat{x}, \tilde{x}, x)$ can be used as local coordinates on $\mathcal{L}$, i.e. we set $y'_I = (\hat{x}, \tilde{x})$. Here we assume that, for given x the map $(\xi/\|\xi\|, e) \mapsto (\hat{x}, \tilde{x}) \in \partial X \times \partial X$ is a local diffeomorphism. Then e is given on $\mathcal{L}$ by a map $(\hat{x}, \tilde{x}, x) \mapsto e(\hat{x}, \tilde{x}, x)$. This map defines e also on an open neighborhood of $\mathcal{L}$ in $T^*Y'\backslash 0$, which by the second part of the proof of Lemma 3.5.1, leads to a choice of symplectic coordinates on a neighborhood of $\mathcal{L}$. For the Maslov-type phase function this choice simply means that $S_{MN}(y'_I, \eta'_J, x, \epsilon)$ does not depend on ϵ.

3.6 Modeling: Joint formulation

In this section we match the expression for the data modeled using the smooth jump (Kirchhoff) approximation to the expressions for the Born modeled data we obtained in Section 3.4. The smooth medium above the interface plays the role of the background medium in the Born approximation.

From Theorem 3.3.1 it follows that reflection of an incident N-mode with covector $\tilde{\xi}_0$ into a scattered M-mode with covector $\hat{\xi}_0$ can take place if the frequencies are equal and $\hat{\xi}_0 + \tilde{\xi}_0$ is normal to the interface. In other words, $\hat{\xi}_0 + \tilde{\xi}_0$ must be in the wavefront set of the singular function of the interface, $\delta(z_n(x))$. Given $\tilde{\xi}_0, \hat{\xi}_0$ one can identify $\mu(M), \nu(N)$, and define (at least to highest order) the reflection coefficient as a function of $(x, \hat{\xi}_0, \tilde{\xi}_0)$, $R^{\text{prin}}_{MN}(x, \hat{\xi}_0, \tilde{\xi}_0) = R^{\text{prin}}_{\mu(M),\nu(N)}(z'(x), \zeta'(\tilde{\xi}_0), \tau)$. This factor can now be viewed as a function of coordinates (y'_I, x, η'_J) or of coordinates (x, ξ, e)

on Λ_{MN} (strictly speaking only defined for x in the interface, and ξ normal to the interface). To highest order it does not depend on $\|\xi\|$ and is it simply a function of (x, e). We obtain the following result, which is a generalization of the Kirchhoff approximation. The normalization factor $\|\frac{\partial z_n}{\partial x}\|$ of the δ-function is such that integral $\int \|\frac{\partial z_n}{\partial x}\| \delta(z_n(x))\,\mathrm{d}x$ is an integral over the surface $z_n = 0$ with Euclidean surface measure in x coordinates.

Theorem 3.6.1 *Suppose Assumptions 1, 2, 3, 4 are satisfied, microlocally for the relevant part of the data. Let $\Phi_{MN}(y', x, \eta'_J), \mathcal{B}_{MN}(y'_I, x, \eta'_J)$ be phase and amplitude as in Theorem 3.4.2, but now for the smooth medium above the interface. The data modeled with the smooth jump model is given microlocally by*

$$G^{\text{refl}}_{MN}(y') = (2\pi)^{-\frac{|J|}{2} - \frac{3n-1-c}{4}} \int (\mathcal{B}_{MN}(y'_I, x, \eta'_J) 2\mathrm{i}\tau(\eta') R_{MN}(y'_I, x, \eta'_J) + \text{l.o.t.}) \\ \times \mathrm{e}^{\mathrm{i}\Phi_{MN}(y', x, \eta'_J)} \left\| \frac{\partial z_n}{\partial x} \right\| \delta(z_n(x))\,\mathrm{d}\eta'_J\,\mathrm{d}x. \tag{3.64}$$

i.e. by a Fourier integral operator with canonical relation Λ_{MN} and order $\frac{n-1+c}{4} - 1$ acting on the distribution $\|\frac{\partial z_n}{\partial x}\| \delta(z_n(x))$.

Proof We write the distribution kernel of the reflected data (3.28) in a form similar to (3.37). First recall the reciprocal expression for the Green's function (3.24),

$$G_N(x(z), \tilde{x}, t_0) = (2\pi)^{-\frac{|\tilde{J}|+1}{2} - \frac{2n+1}{4}} \int \mathcal{A}_N(\tilde{x}_{\tilde{I}}, x(z), \tilde{\xi}_{\tilde{J}}, \tau) \mathrm{e}^{\mathrm{i}\phi_N(\tilde{x}, x(z), t_0, \tilde{\xi}_{\tilde{J}}, \tau)}\,\mathrm{d}\tilde{\xi}_{\tilde{J}}\,\mathrm{d}\tau.$$

By using Theorem 3.3.1, and doing an integration over a t and a τ variable one finds that the Green's function for the reflected part is given by

$$G^{\text{refl}}_{MN}(\hat{x}, \tilde{x}, t) = (2\pi)^{-\frac{|\hat{J}|+|\tilde{J}|+1}{2} - n} \\ \times \int_{z_n=0} \Big(2\mathrm{i}\tau \mathcal{A}_M(\hat{x}_{\hat{I}}, x(z), \hat{\xi}_{\hat{J}}, \tau) \mathcal{A}_N(\tilde{x}_{\tilde{I}}, x(z), \tilde{\xi}_{\tilde{J}}, \tau) R_{\mu(M)\nu(N)}(z, \zeta', \tau) + \text{l.o.t.}\Big) \\ \times \mathrm{e}^{\mathrm{i}\Phi_{MN}(\hat{x}, \tilde{x}, t, x(z), \hat{\xi}_{\hat{J}}, \tilde{\xi}_{\tilde{J}}, \tau)} \left|\det \frac{\partial x}{\partial z}\right| \left\| \frac{\partial z_n}{\partial x} \right\| \mathrm{d}\hat{\xi}_{\hat{J}}\,\mathrm{d}\tilde{\xi}_{\tilde{J}}\,\mathrm{d}\tau\,\mathrm{d}z', \tag{3.65}$$

where ζ' depends on $(x(z), \tilde{\xi}_0)$ (the indices μ, ν for the reflection coefficients have been explained in Section 3.3). The integration $\int \mathrm{d}z'$ is now replaced by $\int \delta(z_n) \mathrm{d}z$. The latter can be transformed back to an integral over x. Thus we obtain

$$(2\pi)^{-\frac{|\hat{J}|+|\tilde{J}|+1}{2} - n} \\ \times \int \Big(2\mathrm{i}\tau \mathcal{A}_M(\hat{x}_{\hat{I}}, x, \hat{\xi}_{\hat{J}}, \tau) \mathcal{A}_N(\tilde{x}_{\tilde{I}}, x, \tilde{\xi}_{\tilde{J}}, \tau) R_{\mu(M)\nu(N)}(z(x), \zeta'(\tilde{\xi}_J, x), \tau) + \text{l.o.t.}\Big) \\ \times \mathrm{e}^{\mathrm{i}\Phi_{MN}(\hat{x}, \tilde{x}, t, x, \hat{\xi}_{\hat{J}}, \tilde{\xi}_{\tilde{J}}, \tau)} \left\| \frac{\partial z_n}{\partial x} \right\| \delta(z_n(x))\,\mathrm{d}\hat{\xi}_{\hat{J}}\,\mathrm{d}\tilde{\xi}_{\tilde{J}}\,\mathrm{d}\tau\,\mathrm{d}x. \tag{3.66}$$

This formula is very similar to (3.37), only the amplitude is different and $\frac{\delta c_{ijkl}(x)}{\rho(x)}, \frac{\delta\rho(x)}{\rho(x)}$ is replaced by the δ-function $\|\frac{\partial z_n}{\partial x}\|\delta(z_n(x))$. Also the factors $w_{MN;ijkl}, w_{MN;0}$ depend only on the background medium, while $R_{\mu(M)\nu(N)}$ depends on the total medium. The phase function Φ_{MN} now comes from the smooth medium above the reflector.

The data is modeled by $G^{\mathrm{refl}}_{MN}(\hat{x},\tilde{x},t)$ with $(\hat{x},\tilde{x},t)$ in the acquisition manifold, as is explained below Lemma 3.4.1. We follow the approach of Section 3.4, and do a coordinate transformation $(\hat{x},\tilde{x},t) \mapsto (y', y'')$, such that the acquisition manifold is given by $y'' = 0$. It follows that under Assumptions 3, 4 the data is the image of a Fourier integral operator acting on $\|\frac{\partial z_n}{\partial x}\|\delta(z_n(x))$ and that it is given by (3.64). □

3.7 Inverse scattering revisited

In this section we present the main results of this chapter. We first construct a Fourier integral operator and a reflectivity function, which is a function of subsurface position and the additional coordinate e. The data is modeled by letting the Fourier integral operator act on the reflectivity. The construction is such that this Fourier integral operator is invertible. We discuss its inverse. Finally a set of pseudodifferential operators is constructed that annihilates the data if the smooth part of the medium above the reflector is correctly chosen.

3.7.1 Invertible transformation into subsurface coordinates

We now construct the reflectivity function and the operator that maps it to seismic data. This is done by applying the results of Section 3.5 to the Kirchhoff modeling formula (3.64), and its equivalent in the Born approximation (3.37).

Theorem 3.7.1 *Suppose microlocally Assumptions 1, 2, 3, 4, 5 are satisfied. Let H_{MN} be the Fourier integral operator with canonical relation given by the extended map $(x,\xi,e,\epsilon) \mapsto (y',\eta')$ constructed in Section 3.5, and with amplitude to highest order given by $(2\pi)^{\frac{n}{2}}(2\mathrm{i}\tau)\mathcal{B}_{MN}(y'_I, x, \eta'_J, \epsilon)$, such that $\mathcal{B}_{MN}(\epsilon = 0)$ is as given in Theorem 3.4.2. Then the data, in both Born and Kirchhoff approximations, is given by H_{MN} acting on a distribution $r_{MN}(x,e)$ of the form*

$$r_{MN}(x,e) = (\text{pseudo})(x, D_x, e)(\text{distribution})(x), \tag{3.67}$$

For the Kirchhoff approximation this distribution equals $\|\frac{\partial z_n}{\partial x}\|\delta(z_n(x))$, while the principal symbol of the pseudodifferential operator equals $R_{MN}(x,e)$, so to highest order $r_{MN}(x,e) = R_{MN}(x,e)\|\frac{\partial z_n}{\partial x}\|\delta(z_n(x))$. For the Born approximation the function $r_{MN}(x,e)$ is given by a pseudodifferential operator acting on $\left(\frac{\delta c_{ijkl}(x)}{\rho(x)}, \frac{\delta\rho(x)}{\rho(x)}\right)_\alpha$, with principal symbol $(2\mathrm{i}\tau(x,\xi,e))^{-1}w_{MN;\alpha}(x,\xi,e)$, see (3.39).

Proof We do the proof for the Kirchhoff approximation using (3.64); for the Born approximation the proof is similar. Since Assumption 5 is satisfied, the projection $\pi_{Y'}$ of Λ_{MN} into $T^*Y'\backslash 0$ is an embedding, and the image is a coisotropic submanifold of $T^*Y'\backslash 0$. Therefore we can apply Lemma 3.5.1. Formula (3.63) implies that the phase factor $\mathrm{e}^{\mathrm{i}\Phi_{MN}}$ can be written in the form

$$\begin{aligned}\mathrm{e}^{\mathrm{i}\Phi_{MN}(y'_I,x,\eta'_J)} &= \mathrm{e}^{\mathrm{i}(S_{MN}(y'_I,x,\eta'_J,0)+\langle y'_J,\eta'_J\rangle)}\\ &= (2\pi)^{-(n-1-c)}\int \mathrm{e}^{\mathrm{i}(S_{MN}(y'_I,x,\eta'_J,\epsilon)+\langle y'_J,\eta'_J\rangle-\langle e,\epsilon\rangle)}\,\mathrm{d}\epsilon\,\mathrm{d}e;\end{aligned}$$

we define

$$\Psi_{MN}(y',x,e,\eta'_J,\epsilon) = S_{MN}(y'_I,x,\eta'_J,\epsilon)+\langle y'_J,\eta'_J\rangle-\langle e,\epsilon\rangle.$$

Thus the number of phase variables is increased by making use of a stationary phase argument. Let $\mathcal{B}_{MN}(y'_I,x,\eta'_J,\epsilon)$ be as described. Then we obtain

$$\begin{aligned}G^{\mathrm{refl}}_{MN}(y') = (2\pi)^{-\frac{|J|+n-1-c}{2}-\frac{2n-1-c}{2}}&\int\left((2\pi)^{\frac{n}{2}}2\mathrm{i}\tau(\eta')\mathcal{B}_{MN}(y'_I,x,\eta'_J,\epsilon)R_{MN}(x,e)+\text{l.o.t.}\right)\\ &\times \mathrm{e}^{\mathrm{i}\Psi_{MN}(y',x,e,\eta'_J,\epsilon)}\left\|\frac{\partial z_n}{\partial x}\right\|\delta(z_n(x))\,\mathrm{d}\eta'_J\,\mathrm{d}\epsilon\,\mathrm{d}x\,\mathrm{d}e. \qquad (3.68)\end{aligned}$$

In this formula the data is represented as a Fourier integral operator acting on $\|\frac{\partial z_n}{\partial x}\|\delta(z_n(x))$ considered as a function of (x,e). Multiplying by H_{MN}^{-1} gives a pseudodifferential operator of the form described acting on $\|\frac{\partial z_n}{\partial x}\|\delta(z_n(x))$. Thus we obtain the result. □

3.7.2 The inversion operator

The operator H_{MN} is invertible. A choice of phase function and amplitude for its inverse is given by (see Chapter 8 of Treves [58])

$$-\Psi_{MN}(y',x,e,\eta'_J,\epsilon),\qquad \mathcal{B}_{MN}(y'_I,x,\eta'_J,\epsilon)^{-1}\left|\det\frac{\partial(y',\eta')}{\partial(y'_I,x,\eta'_J,\epsilon)}\right|,$$

respectively. Thus microlocally an explicit expression for $r_{MN}(x,e)$ in terms of the data is given by

$$\begin{aligned}r_{MN}(x,e) = \int \mathcal{B}_{MN}(y'_I,x,\eta'_J,\epsilon)^{-1}&\left|\det\frac{\partial(y',\eta')}{\partial(y'_I,x,\eta'_J,\epsilon)}\right|\\ &\times \mathrm{e}^{-\mathrm{i}\Psi_{MN}(y',x,e,\eta'_J,\epsilon)}d_{MN}(y')\,\mathrm{d}\eta'_J\,\mathrm{d}\epsilon\,\mathrm{d}y'. \qquad (3.69)\end{aligned}$$

Since the function $r_{MN}(x,e)$ is to highest order equal to the product of reflection coefficient and the singular function of the reflector surface, this reconstruction of the function $r_{MN}(x,e)$ leads to the following result for Kirchhoff data.

Corollary 3.7.2 *Suppose that the medium above the reflector is given, and that it satisfies Assumptions 1, 2, 3, 4, 5. Then one can reconstruct the position of the interface and the angle dependent reflection coefficient $R_{\mu\nu}(x, e)$ on the interface.*

The motivation for Lemma 3.5.1 can be explained in case e is chosen to be the scattering angle/azimuth. Suppose there is high-frequency data that is not from a given model. In the Kirchhoff case this may be because the medium above the interface is not correctly chosen, or because the data cannot be modeled at all by Kirchhoff modeling. To such data there is no natural value of the scattering angle/azimuth associated. So to transform it to (x, e) coordinates the value of e must be chosen. This is precisely the choice that we have in the proof of Lemma 3.5.1, where the function $e(y', \eta')$ on $T^*Y'\backslash 0$ is chosen.

3.7.3 Annihilators of the data

The result of the previous subsections gives information on the problem of reconstructing the smooth background medium (or, in the Kirchhoff approximation, the smooth medium parameters above/in between the interfaces). If $n-1-c > 0$ there is a redundancy in the data through the variable e. If the smooth medium parameters (above the interface) are correct, then applying the operator H_{MN}^{-1} of Theorem 3.7.1 to the data results in a reflectivity function $r_{MN}(x, e)$, such that the position of the singularities does not depend on e. The fact that the inverted data should 'line up' in the variable e can be used as a criterion to assess the accuracy of the background medium.

One way to measure how well the data line up is by taking the derivative with respect to e. If $r_{MN}(x, e)$ depends smoothly on e as in (3.67), then $\frac{\partial}{\partial e} r_{MN}(x, e)$ is one order less singular than if it would not have this smooth dependence on e (for instance a δ function versus its derivative in the Kirchhoff case). Taking also the factor in front of the δ function of r_{MN} into account, see (3.67), we obtain that to the highest two orders

$$\left(R_{MN}(x, e) \frac{\partial}{\partial e} - \frac{\partial R_{MN}^{\mathrm{prin}}}{\partial e}(x, e) \right) r_{MN}(x, e) = 0. \tag{3.70}$$

If $R_{MN}(x, e)$ is nonzero then the lower order terms can be chosen such that this equation is valid to all orders.

Conjugating the differential operator of (3.70) with the invertible Fourier integral operator H_{MN} we obtain a pseudodifferential operator on $\mathcal{D}'(Y')$. Thus we obtain the following corollary of Theorem 3.7.1

Corollary 3.7.3 *Let the pseudodifferential operators $W_{MN}(y', D_{y'})$ be given by*

$$W_{MN}(y', D_{y'}) = H_{MN} \left(R_{MN}(x, e) \frac{\partial}{\partial e} - \frac{\partial R_{MN}}{\partial e}(x, e) \right) H_{MN}^{-1}.$$

Then for Kirchhoff data $d_{MN}(y')$ we have to the highest two orders

$$W_{MN}(y', D_{y'})d_{MN}(y') = 0. \tag{3.71}$$

For values of e where $R_{MN}(x, e) \neq 0$ the operator $W_{MN}(y', D_{y'})$ can be chosen such that (3.71) is valid to all orders.

In principle the operators $W_{MN}(y', D_{y'})$ can be used to obtain a quantitative criterion of how well the data line up. Symes [55] discusses such criteria for acoustic media using the offset coordinate.

Notation

We use the notation $Q(x, D)$ for a pseudodifferential operator with symbol $Q(x, \xi)$, $\mathcal{Q}(x, x_0)$ for its distribution kernel and $Q^{\text{prin}}(x, \xi)$ for its principal symbol.

General

δ_{ij}	Kronecker delta
n	p. 56
x	p. 56
$X \subset \mathbb{R}^n$	p. 56
t	p. 56
$Y', y' \in Y'$	p. 56
$z = (z', z_n)$	p. 70
$e \in E$	p. 81
$\xi, \eta, \zeta, \tau, \epsilon$	cotangent vectors with x, y, z, t, e
$\pi_X, \pi_{Y'}$	p. 80

Subscripts

i, j, k, l	p. 55,60
M, N	p. 60
I, J	p. 65
x_I	p. 65
a	p. 71
μ, ν	p. 71

Field quantities

$\rho(x)$	p. 60
$c_{ijkl}(x)$	p. 60
$\delta c_{ijkl}(x), \delta\rho(x)$	p. 75
$g_\alpha(x)$	p. 80
$u_i(x, t)$	p. 60
$f_i(x, t)$	p. 60
$u_M(x, t), f_M(x, t)$	p. 61
$u_{M,\pm}(x, t), f_{M,\pm}(x, t)$	p. 63
$V_a(x, t)$	p. 71
$V_\mu(x, t)$	p. 72
$d_{MN}(y')$	p. 90
$r_{MN}(x, e)$	p. 89

(Pseudo-)differential operators

P_{il}	p. 60
A_{il}	p. 61
$Q_{iM}(x, D)$	p. 60
$P_M(x, D)$	p. 61
$A_M(x, D)$	p. 61
$B_M(x, D)$	p. 63
$R^0_{\mu\nu}(z, D_{z'}, D_t)$	p. 71
$R_{\mu\nu}(z, D_{z'}, D_t)$	p. 71
$N_{MN;\alpha\beta}(x, D)$	p. 80
$W_{MN}(y', D_{y'})$	p. 91

FIOs and related quantities

$x_M(x_0, \xi_0, t)$, $\xi_M(x_0, \xi_0, t)$	p. 64
$C_{M,\pm}, C_M$	p. 64,65
$\phi_{M,\pm}, \phi_M$	p. 65
$\mathcal{A}_{M,\pm}$	p. 66
$\mathcal{A}_M$	p. 67
G_M	p. 62
$G_{M,\pm}$	p. 63
$\delta G_{il}, \delta G_{MN}$	p. 75
$F_{MN;ijkl}, F_{MN;0}$	p. 78
$F_{MN;\alpha}$	p. 80
Φ_{MN}	p. 76
$\mathcal{B}_{MN}$	p. 76
$w_{MN;ijkl}, w_{MN;0}$	p. 76
$\Lambda_{0,MN}$	p. 76
Λ_{MN}	p. 78
$\mathcal{L}$	p. 80
M_{MN}	p. 86
Ψ_{MN}	p. 87
H_{MN}	p. 89

Chapter 4

Linearized inversion when traveltime injectivity is violated

Abstract

In this chapter we analyze the seismic inverse problem for acoustic media in the presence of multipathing. The problem is to determine the acoustic wavespeed in the interior of the medium from measurements at the boundary. We make two approximations. First a highfrequency approximation. Secondly we perform a linearization or Born approximation where the wavespeed is written as the sum of lowfrequency background $c(x)$ and a highfrequency perturbation $\delta c(x)$. We analyze the problem of reconstructing the highfrequency component given the lowfrequency background.

It is well known that under quite general conditions the linearized forward operator $F : \delta c \mapsto$ data is a Fourier integral operator (FIO). If the so called traveltime injectivity condition is satisfied then the normal operator $N = F^*F$ is an invertible pseudodifferential operator and (asymptotically) a least squares inverse exists.

Our main question is whether F is still invertible if the traveltime injectivity condition is violated. In that case the normal operator is the sum of an invertible pseudodifferential operator and a nonmicrolocal part. We first give conditions when this nonmicrolocal part is an FIO. After that we investigate when the nonmicrolocal part is less singular than the pseudodifferential part, in which case F is still invertible. In the usual imaging method the nonmicrolocal terms would lead to artifacts, for which we obtain estimates. We give an example where the nonmicrolocal part is as singular as the pseudodifferential part of N. We show also that in generic smooth media the nonmicrolocal part is less singular.

4.1 Introduction

In this chapter we study seismic imaging in the presence of multipathing. In particular we study whether it is possible to obtain a correct image in certain situations where the rays do not satisfy the so called traveltime injectivity condition. We first describe some of the ideas in seismic imaging in the presence of multipathing. Our main reference for this topic is Ten Kroode, Smit and Verdel [34]. For elastic media this theory is also described in Section 3.4.

The propagation of sound waves in a medium with soundspeed $c(x)$ and constant density is described by the acoustic wave operator

$$P = \frac{1}{c^2(x)} \frac{\partial^2}{\partial t^2} - \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2}.$$

The acoustic velocity field $u(x,t)$ due to a source $f(x,t)$ satisfying $f(x,t) = 0$ if $t < 0$ is given by the solution to the linear partial differential equation

$$Pu = f$$

that satisfies the initial conditions $u(x,0) = 0, \frac{\partial u}{\partial t}(x,0) = 0$. If the source is a delta function $f(x,t) = \delta(x-s)\delta(t)$, where s is the source position, then the solution is the fundamental solution or Green's function and denoted by $G(x,s,t)$.

In the idealized seismic experiment we assume we have an acoustic medium X, with boundary ∂X. Often the medium is the subsurface described by $\{x \in \mathbb{R}^n \,|\, x_n > 0\}$ with boundary $\partial X = \mathbb{R}^{n-1}$, where $n = 2$ or 3. One measures the signal due to different sources, at a number of receivers during some time interval $I_t =]0, T[$. Sources and receivers are distributed over source and receiver manifolds Σ_s, Σ_r that are assumed to be open parts of the boundary ∂X. The idealized seismic dataset is the set

$$D = \{G(r,s,t) \,|\, r \in \Sigma_r, s \in \Sigma_s, t \in I_t\}.$$

Essentially one is interested in determining the soundspeed $c(x)$ in the medium from the data.

To model and invert the data we make two important approximations. The first is a highfrequency approximation. This allows us to use the techniques of microlocal analysis. The main idea of microlocal analysis is that the singularities or highfrequency part of a function $f(x)$ can be localized both with respect to position x as with respect to wavevector ξ, which is an element of the cotangent space $T_x^* X \backslash 0$. In particular this allows one to resolve the multipathing.

The second approximation is that we do a linearization or Born approximation in the medium coefficients. The full soundspeed is written as the sum of a smooth (lowfrequency) background velocity c and a perturbation δc, that contains the singularities (highfrequency part). The Green's function is written as $G + \delta G$, where

$$\delta G(r,s,t) = -\int G(r,x,t-t')\,\delta P\, G(x,s,t')\,\mathrm{d}x\,\mathrm{d}t'. \tag{4.1}$$

The map $F : \delta c(x) \mapsto \delta G(r,s,t)$ is called the linearized forward map. We assume that, after preprocessing, the data contains only singly reflected waves (no direct

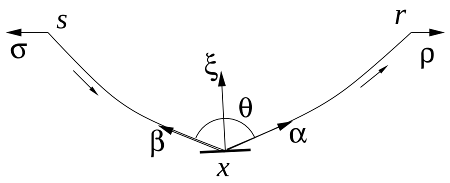


Figure 4.1: Picture of the scattering and the different coordinates

waves and no multiple reflection), which implies that δG describes all the data. The problem that we address here is the reconstruction of the singular part $\delta c(x)$, given the smooth background velocity $c(x)$.

Under quite general conditions (see Ten Kroode, Smit and Verdel [34]) the linearized forward operator F is a Fourier integral operator (FIO), that is, it has a kernel of the form

$$F(r,s,t,x) = \int A(r,s,t,x,\Theta)\mathrm{e}^{\mathrm{i}\Phi(r,s,t,x,\Theta)}\,\mathrm{d}\Theta.$$

(see Duistermaat [17], Hörmander [27], Treves [58]). This was proved by Rakesh [46], Hansen [22], while in the present setting it was proved by Ten Kroode, Smit and Verdel [34]. For elastic media this result is given in Theorem 3.4.2. Let $Y' = \Sigma_r \times \Sigma_s \times I_t$. The positions in $T^*Y'\backslash 0$ where F maps the singularities of δc are given by the canonical relation $\Lambda \subset T^*Y'\backslash 0 \times T^*X\backslash 0$. It can be parametrized conveniently by position of the scattering point x, frequency τ and the phase directions of the incoming and outgoing rays at the scattering point α and β, that were denoted by $\hat{\alpha}$ and $\tilde{\alpha}$ respectively in the previous chapter (see below equation (3.47)). Using these parameters we have

$$\begin{aligned}\Lambda = \{&(r(x,\alpha), s(x,\beta), \phi_r(x,\alpha)+\phi_s(x,\beta), \rho(x,\alpha,\tau), \sigma(x,\beta,\tau), \tau,\\ &x, -\frac{\tau}{c(x)}(\alpha+\beta)) \,|\, x\in X, \tau\in\mathbb{R}, \alpha,\beta\in S^{n-1}\}.\end{aligned} \tag{4.2}$$

Here $r(x,\alpha), s(x,\beta)$ are the receiver and source coordinates respectively, in other words the positions where the (x,α) or (x,β) ray hits the receiver or source surface. By $\rho(x,\alpha,\tau), \sigma(x,\beta,\tau)$ we denote the corresponding cotangent vectors, and ϕ_r, ϕ_s are the traveltimes from x to the receiver and the source respectively. Sometimes we will also use the tangential slowness vectors $p_r = -\tau^{-1}\rho, p_s = -\tau^{-1}\sigma \in \mathbb{R}^{n-1}$ (Note that the notation differs from the notation in [34, 52]). A picture of the scattering and the different coordinates is given in Figure 4.1. It describes the scattering due to a singular component of δc with coordinates $x, \xi = -\frac{(\alpha+\beta)\tau}{c(x)}$. A highfrequency component with coordinates s,σ leads to a signal that propagates along a uniquely determined ray. At x it interacts with δc (the multiplication in (4.1)), leading to a scattered signal that propagates along a different ray to the receiver.

To find an inverse for F we use the method of least squares, since there is a redundancy in the data. The operator F has a left inverse if and only if the normal operator $N = F^*F$ is invertible. In that case a left inverse is given by

$$N^{-1}F^*. \tag{4.3}$$

The operator N plays therefore an important role. Beylkin [8] has shown that if there are no caustics on the rays connecting source and receiver points to the scattering points, then the normal operator is an invertible pseudodifferential operator of order $n-1$. Ten Kroode e.a. [34] have shown that this result still holds when the no-caustics assumption is replaced by the less restrictive traveltime injectivity condition (part ii) of Assumption 5, we will say more about this assumption in Section 4.3 below). The medium satisfies the traveltime injectivity condition if, given a ray defined by receiver coordinate and slowness and another ray defined by source coordinate and slowness and the travel time, one can uniquely solve for the intersection point if the rays intersect. A pseudodifferential operator of order $n-1$ is a continuous map $H_{(s)} \to H_{(s-(n-1))}$. It can be inverted asymptotically, provided that the amplitude is nonzero, so in this case the operator F has an asymptotic left inverse. See also Nolan and Symes [41], who discuss the case of lower dimensional acquisition manifolds, and Section 3.4 of this thesis, where the case of elastic media is discussed for general acquisition manifolds.

So if the traveltime injectivity condition is satisfied there is an inverse for the linearized forward map F. What if it is not satisfied? This is the subject of the present chapter. It has been discussed by Ten Kroode e.a. [34]. However, there is more to say about this subject, and we show that one of their results (Theorem 3.2) is actually incorrect.

Examples of rayconfigurations that cause this situation are given in Figure 4.2 in Section 4.3. In this case one cannot distinguish between scattering at x and $\bar{x}$ using data for this value of the scattering angle θ. Often by using data related to different values of θ one can still make the distinction, so that inversion is still possible. In some cases this is not possible.

We will analyze the problem of inversion in such situations using the normal operator. Ten Kroode e.a. [34] showed that the normal operator is the sum of a pseudodifferential part $N_{\psi\text{do}}$ and a nonmicrolocal part N_{nml} (we call it the nonmicrolocal part and not the nonlocal part, because nonmicrolocal refers to the highfrequency part, which is nonlocal on the cotangent bundle $T^*X\backslash 0$). If the nonmicrolocal part is less singular than the pseudodifferential part then N can still be inverted asymptotically, treating N_{nml} as a perturbation. This approach also gives an idea of the artifact (error) that arises when we use $N_{\psi\text{do}}^{-1}$ instead of N^{-1} in (4.3), i.e. when we ignore the nonmicrolocal part.

We give an overview of the rest of the chapter. The first step is to check when the normal operator is a Fourier integral operator. This is the subject of Section 4.3, using some of preliminaries that are explained in Section 4.2. We simplify a condition obtained by Ten Kroode e.a. [34] for the nonmicrolocal part to be a Fourier integral operator of order $\frac{n-1}{2}$.

Next we note that this does not imply that N_{nml} is continuous $H_{(s)} \to H_{(\frac{n-1}{2})}$. The question between which Sobolev space N_{nml} is continuous is in general quite difficult. It involves the study of so called degenerate Fourier integral operators (see e.g. Hörmander [27], Seeger [47, 48], Phong and Stein [43, 44]). We discuss this in Section 4.4. The nonmicrolocal part is at most as singular as the pseudodifferential part. We give conditions when it is less singular than the pseudodifferential part. Using results of Seeger [47] we derive detailed estimates for the case $n = 2$.

In Section 4.5 we start with an example of this singular behavior. We also give an example where the nonmicrolocal part is less singular than the pseudodifferential part and an example where the condition of Section 4.3 is violated.

In Section 4.6 we show that in generic smooth media (i.e. media with no special properties such as symmetries) the nonmicrolocal part is less singular than the pseudodifferential part. We use results on the perturbation of rays when the medium is perturbed, that are given in Appendix 4.A. We end with a discussion of the results.

4.2 Preliminaries, the linearized forward operator

In this section the construction of rays, traveltimes and of the derivatives of rays is discussed. After that we explain precisely the statement that F is an FIO.

The rays can be found by solving a Hamilton system in T^*X, with parameter t. The Hamiltonian is given by $H(x,\xi) = c(x)\|\xi\|$, the system is

$$\begin{aligned}\frac{\partial x_i}{\partial t} &= \frac{\partial H}{\partial \xi_i}(x,\xi) = c(x)\frac{\xi_i}{\|\xi\|},\\ \frac{\partial \xi_i}{\partial t} &= -\frac{\partial H}{\partial x_i}(x,\xi) = -\frac{\partial c}{\partial x_i}(x)\|\xi\|.\end{aligned} \tag{4.4}$$

The mapping $(x_0,\xi_0,t) \mapsto (x(x_0,\xi_0,t),\xi(x_0,\xi_0,t))$, that maps initial values x_0,ξ_0 to the solution of (4.4) at time t is called the Hamilton flow. The Hamilton flow is homogeneous in ξ_0,ξ and it conserves the Hamiltonian. If we let $\alpha_0 = \frac{\xi_0}{\|\xi_0\|}$, $\alpha = \frac{\xi}{\|\xi\|}$, then the Hamilton flow gives a map $(x_0,\alpha_0,t) \mapsto (x(x_0,\alpha_0,t),\alpha(x_0,\alpha_0,t))$.

The Jacobi matrix $\frac{\partial(x,\xi)}{\partial(x_0,\xi_0)}(t)$ satisfies an ordinary differential equation along the ray. By differentiating (4.4) we obtain the Jacobi or neighboring ray equations

$$\begin{aligned}\frac{\partial}{\partial t}\frac{\partial x_i}{\partial(x_0,\xi_0)} &= \sum_j\left(\frac{\partial^2 H}{\partial \xi_i \partial x_j}\frac{\partial x_j}{\partial(x_0,\xi_0)} + \frac{\partial^2 H}{\partial \xi_i \partial \xi_j}\frac{\partial \xi_j}{\partial(x_0,\xi_0)}\right),\\ \frac{\partial}{\partial t}\frac{\partial \xi_i}{\partial(x_0,\xi_0)} &= \sum_j\left(-\frac{\partial^2 H}{\partial x_i \partial x_j}\frac{\partial x_j}{\partial(x_0,\xi_0)} - \frac{\partial^2 H}{\partial x_i \partial \xi_j}\frac{\partial \xi_j}{\partial(x_0,\xi_0)}\right).\end{aligned} \tag{4.5}$$

The Jacobi matrix is symplectic.

We will now investigate the map $(x_0,\alpha_0) \mapsto (r,p_r,\phi_r)$, that maps a pair (x_0,α_0) in the medium to receiver coordinate r, receiver slowness p_r and traveltime ϕ_r. For the map $(x_0, \beta_0) \mapsto (s, p_s, \phi_s)$ the same results are valid. If $(x_0, \alpha_0, t) \mapsto (x, \alpha)$ denotes the Hamilton flow to a neighborhood of the receiver point, and x_n is the coordinate normal to the surface, then this map is obtained by solving the traveltime to the receiver $\phi_r(x_0, \alpha_0)$ from

$$x_n(x_0, \alpha_0, \phi_r) = 0$$

and then setting

$$\begin{aligned} r(x_0, \alpha_0) &= (x_1(x_0, \alpha_0, \phi_r(x_0, \alpha_0)), \dots , x_{n-1}(x_0, \alpha_0, \phi_r(x_0, \alpha_0))), \\ p_r(x_0, \alpha_0) &= \frac{1}{c(x(x_0, \alpha_0))}(\alpha_1(x_0, \alpha_0, \phi_r(x_0, \alpha_0)), \dots , \alpha_{n-1}(x_0, \alpha_0, \phi_r(x_0, \alpha_0))). \end{aligned}$$

In this way we can define ϕ_r provided there are no rays that come in tangent to the surface (so called grazing rays). The derivatives of this map are

$$\begin{aligned} \frac{\partial \phi_r}{\partial (x_0, \alpha_0)} &= - \left(\frac{\partial x_n}{\partial t} \right)^{-1} \cdot \frac{\partial x_n}{\partial (x_0, \alpha_0)} \\ \frac{\partial r}{\partial (x_0, \alpha_0)} &= \frac{\partial (x_1, \dots , x_{n-1})}{\partial (x_0, \alpha_0)} - \frac{\partial (x_1, \dots , x_{n-1})}{\partial t} \left(\frac{\partial x_n}{\partial t} \right)^{-1} \frac{\partial x_n}{\partial (x_0, \alpha_0)}. \\ \frac{\partial p_r}{\partial (x_0, \alpha_0)} &= \frac{\partial}{\partial (x_0, \alpha_0)} \left(\frac{\alpha_1}{c(x)}, \dots , \frac{\alpha_{n-1}}{c(x)} \right) \\ &\quad - \left(\frac{\partial}{\partial t} \frac{\alpha_1}{c(x)}, \dots , \frac{\partial}{\partial t} \frac{\alpha_{n-1}}{c(x)} \right) \left(\frac{\partial x_n}{\partial t} \right)^{-1} \frac{\partial x_n}{\partial (x_0, \alpha_0)}. \end{aligned} \tag{4.6}$$

It is not difficult to show that $\frac{\partial (r, p_r, \phi_r)}{\partial (x_0, \alpha_0)}$ is invertible, so that the map $(x_0, \alpha_0) \mapsto (r, p_r, \phi_r)$ is a diffeomorphism. One more property of $\frac{\partial (r, p_r, \phi_r)}{\partial (x_0, \alpha_0)}$ will be needed. The derivative of (r, p_r) with respect to x along the ray vanishes

$$\sum_i \frac{\partial (r, p_r)}{\partial x_{0,i}} \alpha_{0,i} = 0, \tag{4.7}$$

while the derivative of the traveltime ϕ_r along the ray satisfies

$$\sum_i \frac{\partial \phi_r}{\partial x_{0,i}} \alpha_{0,i} = -\frac{1}{c(x_0)}. \tag{4.8}$$

We discuss the linearized forward operator F. Ten Kroode e.a. have shown that under certain conditions it is an FIO. This property is essentially microlocal. Suppose a ray determined by s, p_s and a ray determined by r, p_r intersect in some point x, suppose the rays intersect source resp. receiver surface transversally (no grazing rays) and the scattering angle $\neq \pi$. This leads to a contribution to F that is microlocally an FIO. Since the two rays have at most a finite number of intersections with the same traveltime $t_r + t_s$ there is for a small enough neighborhood of $(r, s, t_r + t_s, \rho, \sigma, \tau)$ in $T^*Y'\backslash 0$ a finite number of such contributions.

Recall that we define $Y' = \Sigma_r \times \Sigma_s \times I_t$, let $y' = (r, s, t) \in Y'$ and denote the corresponding cotangent vector by η'. Let $K \subset T^*Y'\backslash 0$ be the set of points where one of the rays is tangent or scattering over π occurs. This set is closed, so F is microlocally an FIO on an open set $(T^*Y'\backslash 0)\backslash K$. The linearized forward operator F can be made into a globally defined Fourier integral operator by applying a pseudodifferential cutoff $\psi(y', D_{y'})$. Here $\psi(y', \eta')$ is a homogeneous symbol of order 0 that vanishes outside some compact subset $V_1 \subset (T^*Y'\backslash 0)\backslash K$, and is equal to 1 on a compact subset $V_2 \subset V_1$. The operator $\psi(y', D_{y'}) \circ F$ is now a finite sum of local Fourier integral operators.

Another way of dealing with scattering over π and grazing rays is to assume that it is absent, in other words that the set K is empty. This is done by Ten Kroode e.a. [34] and Nolan and Symes [41]. In this case we have to be careful at the boundary of $T^*Y'\backslash 0$. Although microlocally F is an FIO, the microlocal contributions do not satisfy uniform estimates, and therefore they may diverge at the boundary. Strictly speaking this situation should be taken into account in [34, 41]. It is for instance not clear that the symbol of the normal operator (expression (64) in [34]) is finite.

4.3 Conditions for the normal operator to be a FIO

The condition under which the normal operator is a Fourier integral operator were derived by Ten Kroode e.a. [34]. It turns out that these conditions can be simplified and have a geometrical interpretation. We also give a characterization of the canonical relation of this Fourier integral operator. This is important since the continuity of the normal operator between Sobolev spaces depends on the properties of the canonical relation.

The adjoint F^* is a Fourier integral operator of the same order as F with canonical relation

$$\Lambda^* = \{(x, \xi; y', \eta') \,|\, (y', \eta'; x, \xi) \in \Lambda\}.$$

We discuss the composition of canonical relations $\Lambda^* \circ \Lambda$. Let $L = \Lambda^* \times \Lambda$, $M = T^*X\backslash 0 \times \operatorname{diag}(T^*Y'\backslash 0) \times T^*X\backslash 0$. The compose $\Lambda^* \circ \Lambda$ is given by the projection of $L \cap M$ on $T^*X\backslash 0 \times T^*X\backslash 0$, in other words

$$\Lambda^* \circ \Lambda = \{(\bar{x}, \bar{\xi}; x, \xi) \,|\, \text{there is } (y', \eta') \text{ with } (y', \eta'; x, \xi) \in \Lambda, (y', \eta'; \bar{x}, \bar{\xi}) \in \Lambda\}.$$

If L and M intersect cleanly (this will be explained below), then the compose is again a canonical relation, see e.g. Hörmander [26], Theorem 21.2.14, or Treves [58] paragraph 8.5.

In the present case the condition that L and M intersect is that there are $x, \alpha, \beta, \tau, \bar{x}, \bar{\alpha}, \bar{\beta}, \bar{\tau}$, such that

$$\begin{aligned} y'(x, \alpha, \beta, \tau) &= y'(\bar{x}, \bar{\alpha}, \bar{\beta}, \bar{\tau}) \\ \eta'(x, \alpha, \beta, \tau) &= \eta'(\bar{x}, \bar{\alpha}, \bar{\beta}, \bar{\tau}). \end{aligned} \tag{4.9}$$

These equations say that

1. Both $x, \bar{x}$ are on the ray determined by r, p_r and $\alpha, \bar{\alpha}$ are the directions of the ray at x resp. $\bar{x}$.
2. Both $x, \bar{x}$ are on the ray determined by s, p_s and $\beta, \bar{\beta}$ are the directions of the ray at x resp. $\bar{x}$.
3. If $x \neq \bar{x}$ then the equality $\phi_r(x, \alpha) + \phi_s(x, \beta) = \phi_r(\bar{x}, \bar{\alpha}) + \phi_s(\bar{x}, \bar{\beta})$ implies that the (r, p_r)-ray hits x first and then $\bar{x}$, while the (s, p_s) ray first hits $\bar{x}$ and then x. The traveltimes from x to $\bar{x}$ along the two rays are equal.
4. $\tau = \bar{\tau}$.

In other words this means that there are two rays originating in $\bar{x}$, in the directions $\bar{\alpha}$, resp. $-\bar{\beta}$, that intersect in x at the same traveltime. Therefore (4.9) is satisfied if and only if there is t such that

$$\Delta x(\bar{x}, \bar{\alpha}, \bar{\beta}, t) := x(\bar{x}, \bar{\alpha}, t) - x(\bar{x}, \bar{\beta}, -t) = 0, \tag{4.10}$$

and

$$\begin{aligned} x &= x(\bar{x}, \bar{\alpha}, t), \\ \alpha &= \alpha(\bar{x}, \bar{\alpha}, t), \\ \beta &= \beta(\bar{x}, \bar{\beta}, -t). \end{aligned} \tag{4.11}$$

The compose $\Lambda^* \circ \Lambda$ is then given by

$$\{(\bar{x}, \xi(\bar{x}, \bar{\alpha}, \bar{\beta}, t, \bar{\tau})), (x, \xi(x, \alpha, \beta, t, \tau)) \,|\, \text{equations (4.10), (4.11) are satisfied}\}. \tag{4.12}$$

There are two types of solutions to these equations, that we will call solutions of type 1 resp. type 2. First there is the solution where $t = 0$ and $(x, \alpha, \beta) = (\bar{x}, \bar{\alpha}, \bar{\beta})$. Secondly there may be solutions such that $t \neq 0$ and in general $(x, \alpha, \beta) \neq (\bar{x}, \bar{\alpha}, \bar{\beta})$. These give a contribution to the normal operator that is in general not microlocal (i.e. $(x, \xi) \neq (\bar{x}, \bar{\xi})$).

Such a configuration is quite special, i.e. most rays that intersect each other do not give such a situation, because the traveltimes are different (see Figure 4.2a). In Figure 4.2b,c we give two situations that have a nonmicrolocal term. An example of a nonmicrolocal term with no special properties is given in Figure 4.2b. An example where $x = \bar{x}$ is given in Figure 4.2c. In general in the situation of Figure 4.2c one

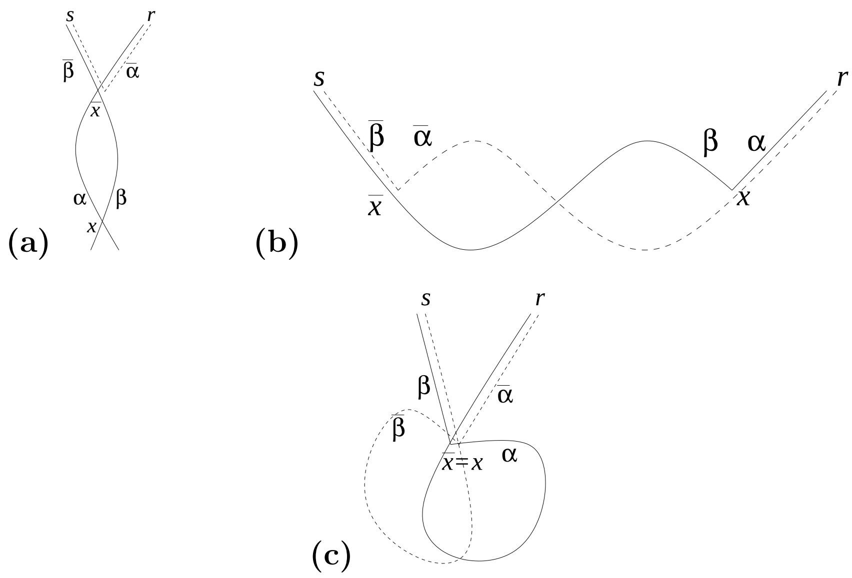


Figure 4.2: Ray configuarations where the Traveltime Injectivity Condition holds (a), is violated in a generic way (b), resp. in a special way such that $x = \bar{x}$ (c).

has $\xi \neq \bar{\xi}$, but we may even have $\frac{\alpha+\beta}{\|\alpha+\beta\|} = \frac{\bar{\alpha}+\bar{\beta}}{\|\bar{\alpha}+\bar{\beta}\|}$, so that $\xi/\|\xi\| = \bar{\xi}/\|\bar{\xi}\|$, or (in $n = 3$ dimensions) $\alpha + \beta = \bar{\alpha} + \bar{\beta}$, so that $\xi = \bar{\xi}$, i.e. a selfintersection of Λ.

The intersection of the manifolds L and M defined above is called clean if

$$\begin{aligned} &L \cap M \text{ is a manifold, and for each element of } L \cap M \\ &TL \cap TM = T(L \cap M). \end{aligned} \tag{4.13}$$

This is equivalent to the statement that L and M intersect transversally in a submanifold of $T^*X\backslash 0 \times T^*Y'\backslash 0 \times T^*Y'\backslash 0 \times T^*X\backslash 0$. For solutions of type 1 we show below that this is automatically the case. For solutions of type 2 we require that the intersection is transversal (clean with excess 0). This means that the matrix

$$M = \begin{pmatrix} \frac{\partial r}{\partial x} & \frac{\partial r}{\partial \alpha} & 0 & \frac{\partial r}{\partial \bar{x}} & \frac{\partial r}{\partial \bar{\alpha}} & 0 \\ \frac{\partial p_r}{\partial x} & \frac{\partial p_r}{\partial \alpha} & 0 & \frac{\partial p_r}{\partial \bar{x}} & \frac{\partial p_r}{\partial \bar{\alpha}} & 0 \\ \frac{\partial s}{\partial x} & 0 & \frac{\partial s}{\partial \beta} & \frac{\partial s}{\partial \bar{x}} & 0 & \frac{\partial s}{\partial \bar{\beta}} \\ \frac{\partial p_s}{\partial x} & 0 & \frac{\partial p_s}{\partial \beta} & \frac{\partial p_s}{\partial \bar{x}} & 0 & \frac{\partial p_s}{\partial \bar{\beta}} \\ \frac{\partial \phi}{\partial x} & \frac{\partial \phi}{\partial \alpha} & \frac{\partial \phi}{\partial \beta} & \frac{\partial \phi}{\partial \bar{x}} & \frac{\partial \phi}{\partial \bar{\alpha}} & \frac{\partial \phi}{\partial \bar{\beta}} \end{pmatrix}.$$

coming for linearizing (4.9) has maximal rank (we omit the part relating to $\tau, \bar{\tau}$ since it is trivial). In the following lemma we show that in fact it is sufficient to consider the much smaller matrix

$$C = \begin{pmatrix} \frac{\partial x(\bar{x},\bar{\alpha},t)}{\partial \bar{x}} - \frac{\partial x(\bar{x},\bar{\beta},-t)}{\partial \bar{x}} & \frac{\partial x(\bar{x},\bar{\alpha},t)}{\partial \bar{\alpha}} & -\frac{\partial x(\bar{x},\bar{\beta},-t)}{\partial \bar{\beta}} & \frac{\partial x(\bar{x},\bar{\alpha},t)}{\partial t} + \frac{\partial x}{\partial t}(\bar{x},\bar{\beta},-t) \end{pmatrix}. \tag{4.14}$$

coming from linearizing (4.10). Note that this parallels the simplification of the system (4.9) to (4.10). It implies that the relation of the normal operator (4.12) is a manifold, as of course it should be.

Lemma 4.3.1 *Suppose $\bar{x}, \bar{\alpha}, \bar{\beta}, x, \alpha, \beta, t$ are such that equations (4.10) and (4.11) are satisfied. Then the rank of the matrix M satisfies*

$$\operatorname{rank} M = 3n - 3 + \operatorname{rank} C. \tag{4.15}$$

In particular $\operatorname{rank} M$ *is maximal if and only if* $\operatorname{rank} C$ *is maximal.*

Proof Let the assumption of the lemma be satisfied. We write the matrix M as a product of two factors, where one factor contains the derivatives $\frac{\partial(r,p_r,\phi_r)}{\partial(x,\alpha)}$, $\frac{\partial(s,p_s,\phi_s)}{\partial(x,\beta)}$ and the other factor relates (x,α,β) to $(\bar{x},\bar{\alpha},\bar{\beta})$. Here we use that the map

$$(\bar{x},\bar{\alpha}) \mapsto (r(\bar{x},\bar{\alpha}), p_r(\bar{x},\bar{\alpha}), \phi_r(\bar{x},\bar{\alpha}))$$

equals the composition of the maps

$$\begin{aligned}
&(\bar{x},\bar{\alpha}) \mapsto (x(\bar{x},\bar{\alpha},t), \alpha(\bar{x},\bar{\alpha},t)),\\
&(x,\alpha) \mapsto (r(x,\alpha), p_r(x,\alpha), \phi_r(x,\alpha)+t).
\end{aligned}$$

The first map is the transport of $(\bar{x},\bar{\alpha})$ over time t using the Hamilton system. This time t needs to be added to the traveltime $\phi(x(\bar{x},\bar{\alpha},t),\alpha(\bar{x},\bar{\alpha},t))$ to obtain the traveltime from $(\bar{x},\bar{\alpha})$ to the surface. The map $(\bar{x},\bar{\beta}) \mapsto (s(\bar{x},\bar{\beta}), p_s(\bar{x},\bar{\beta}), \phi_s(\bar{x},\bar{\beta}))$ can be decomposed in a similar, where the role of the variables $(\alpha,\bar{\alpha},r,p_r)$ and $(\beta,\bar{\beta},s,p_s)$ is interchanged and $t \leftrightarrow -t$. It follows that

$$\begin{aligned}
\frac{\partial(r,p_r,\phi_r)}{\partial(\bar{x},\bar{\alpha})}(\bar{x},\bar{\alpha}) &= \frac{\partial(r,p_r,\phi_r)}{\partial(x,\alpha)}(x,\alpha)\cdot\frac{\partial(x,\alpha)}{\partial(\bar{x},\bar{\alpha})}(\bar{x},\bar{\alpha},t),\\
\frac{\partial(s,p_s,\phi_s)}{\partial(\bar{x},\bar{\beta})}(\bar{x},\bar{\beta}) &= \frac{\partial(s,p_s,\phi_s)}{\partial(x,\beta)}(x,\beta)\cdot\frac{\partial(x,\beta)}{\partial(\bar{x},\bar{\beta})}(\bar{x},\bar{\beta},-t).
\end{aligned}$$

This can also be seen from expression (4.6) above. Thus the matrix M can be decomposed as follows

$$M = A \cdot M',$$

where A and M' are $(4n-3)\times(4n-2)$ resp. $(4n-2)\times(6n-4)$ matrices given by

$$A = \begin{pmatrix}
\frac{\partial(r,p_r)}{\partial(x,\alpha)}(x,\alpha) & \\
 & \frac{\partial(s,p_s)}{\partial(x,\beta)}(x,\beta)\\
\frac{\partial\phi_r}{\partial(x,\alpha)}(x,\alpha) & \frac{\partial\phi_s}{\partial(x,\beta)}(x,\beta)
\end{pmatrix},$$

$$M' = \begin{pmatrix}
I_n & & & \frac{\partial x}{\partial\bar{x}}(\bar{x},\bar{\alpha},t) & \frac{\partial x}{\partial\bar{\alpha}}(\bar{x},\bar{\alpha},t) & \\
 & I_{n-1} & & \frac{\partial\alpha}{\partial\bar{x}}(\bar{x},\bar{\alpha},t) & \frac{\partial\alpha}{\partial\bar{\alpha}}(\bar{x},\bar{\alpha},t) & \\
I_n & & & \frac{\partial x}{\partial\bar{x}}(\bar{x},\bar{\beta},-t) & & \frac{\partial x}{\partial\bar{\beta}}(\bar{x},\bar{\beta},-t)\\
 & & I_{n-1} & \frac{\partial\beta}{\partial\bar{x}}(\bar{x},\bar{\beta},-t) & & \frac{\partial\beta}{\partial\bar{\beta}}(\bar{x},\bar{\beta},-t)
\end{pmatrix}. \tag{4.16}$$

The matrix A has maximal rank since the matrices

$$\frac{\partial(r,p_r,\phi_r)}{\partial(x,\alpha)}(x,\alpha), \qquad \frac{\partial(s,p_s,\phi_s)}{\partial(x,\beta)}(x,\beta)$$

are invertible. But it has a nonzero kernel. To find this kernel we note that

$$\frac{\partial(r,p_r)}{\partial(x,\alpha)}(x,\alpha)\cdot\begin{pmatrix} c(x)\,\alpha \\ 0\end{pmatrix} = 0, \qquad \frac{\partial\phi_r}{\partial(x,\alpha)}(x,\alpha)\cdot\begin{pmatrix} c(x)\,\alpha \\ 0\end{pmatrix} = -1.$$

The same is true with r, p_r, α and s, p_s, β interchanged. Hence

$$\ker A = \operatorname{span}(c(x)\,\alpha, 0, -c(x)\,\beta, 0).$$

Now basic linear algebra says that

$$\operatorname{rank} M = \operatorname{rank} M' - \dim(\operatorname{range} M' \cap \ker A). \tag{4.17}$$

By trying to solve w in the system $v = M'\cdot w$ one finds that $v = (v_{x,1}, v_\alpha, v_{x,2}, v_\beta) \in \operatorname{range} M'$ if and only if $v_{x,1} - v_{x,2} \in \operatorname{range} C'$, where

$$C' = \begin{pmatrix} \frac{\partial x(\bar{x},\bar{\alpha},t)}{\partial \bar{x}} - \frac{\partial x(\bar{x},\bar{\beta},-t)}{\partial \bar{x}} & \frac{\partial x(\bar{x},\bar{\alpha},t)}{\partial \bar{\alpha}} & -\frac{\partial x(\bar{x},\bar{\beta},-t)}{\partial \bar{\beta}} \end{pmatrix}.$$

Hence $\operatorname{rank} M' = 3n - 2 + \operatorname{rank} C'$ and

$$\dim(\operatorname{range} M' \cap \ker A) = \dim(\operatorname{range} C' \cap \operatorname{span}(c(x)\alpha + c(x)\beta)).$$

Using (4.17) and the fact that

$$\operatorname{rank} C = \operatorname{rank} C' + 1 - \dim(\operatorname{range} C' \cap \operatorname{span}(c(x)\alpha + c(x)\beta))$$

the result (4.15) follows. $\square$

We summarize the properties of the intersection of L and M in the following lemma. We use the notation $p = (x, \alpha, \beta, \tau)$. To parametrize Λ, the parameter p varies in some open subset $P \subset X \times S^{n-1} \times S^{n-1} \times \mathbb{R}$. It is important to understand that Λ is an immersed manifold, it is the image of an immersion. So although it may have selfintersections the intersecting parts will be disjoint in the preimage.

Lemma 4.3.2 *The set $L \cap M$ is a union of a microlocal part*

$$\{(x,\xi,y',\eta',y',\eta',x,\xi) \mid (y',\eta';x,\xi) \in \Lambda\} \tag{4.18}$$

and a nonmicrolocal part

$$\{(x(\bar{p}),\xi(\bar{p}),y'(\bar{p}),\eta'(\bar{p}),y'(p),\eta'(p),x(p),\xi(p)) \mid (y'(\bar{p}),\eta'(\bar{p})) = (y'(p),\eta'(p)), \bar{p} \neq p\}. \tag{4.19}$$

For the microlocal part the intersection is clean with excess $n-1$. For the nonmicrolocal part the intersection is transversal if and only if $\operatorname{rank} M$ *is maximal, if and only if* $\operatorname{rank} C$ *is maximal. If $\Lambda_{\mathrm{c}} \subset \Lambda$ is a conically compact subset of Λ, and we replace L by $L_{\mathrm{c}} = \Lambda_{\mathrm{c}}^* \times \Lambda_{\mathrm{c}}$, then the microlocal and the nonmicrolocal part of $L_{\mathrm{c}} \cap M$ are both conically compact.*

Proof The projection $\pi_{Y'}$ of Λ on $T^*Y'\backslash 0$ is immersive. So for any point (x, ξ, y', η') $\in \Lambda$ there is a small open neighborhood $\Lambda_0 \subset \Lambda$ such that $\pi_{Y'}$ acting on Λ_0 is an embedding, and hence $\Lambda_0^* \times \Lambda_0 \cap M = \{(x, \xi, y', \eta', y', \eta', x, \xi) \,|\, (y', \eta', x, \xi) \in \Lambda_0\}$. Since (y', η', x, ξ) is arbitrary we can find subsets $\Lambda_i \subset \Lambda$, with i in some index set I such that $\pi_{Y'}$ acting on Λ_i is an embedding. Let

$$L_{\rm ml} = \bigcup_{i \in I} \Lambda_i^* \times \Lambda_i,$$

then $L_{\rm ml} \cap M$ consists precisely of the solutions of (4.9) of type 1. For each i the intersection $\Lambda_i^* \times \Lambda_i \cap M$ is transversal in the submanifold $T^*X\backslash 0 \times (\operatorname{diag} \pi_{Y'}(\Lambda_i)) \times T^*X\backslash 0$, so the intersection is clean with excess given by $\operatorname{codim} \pi_{Y'}(\Lambda_i) = \dim Y' - \dim X = n - 1$.

Let $L_{\rm nml} = L\backslash L_{\rm ml}$. Clearly the elements of $L_{\rm nml} \cap M$ precisely correspond to the solutions to (4.9) of type 2. By definition the intersection is transversal if and only if $\operatorname{rank} M$ is maximal which by Lemma 4.3.1 is equivalent to $\operatorname{rank} C$ is maximal.

If $\Lambda_{\rm c}$ is a compact subset of Λ, then $L_{\rm c}$ is closed and we may replace M by some conically compact subset of $M_{\rm c} \subset M$. From the formula (4.18) it follows that $L_{\rm c;ml} \cap M_{\rm c}$ is conically compact. Also $L_{\rm c;nml} = L_{\rm c}\backslash L_{\rm c;ml}$ is conically compact, and hence $L_{\rm c;nml} \cap M_{\rm c}$ is conically compact. □

From the lemma it follows that in parameterspace P the microlocal and nonmicrolocal terms are well separated. So if $\Lambda_{\rm c}$ is some compact subset of Λ we can choose a finite number of Λ_i, such that $\Lambda_{\rm c} \subset \bigcup_i \Lambda_i$ and the compose $\Lambda_i^* \circ \Lambda_j$ is either fully diagonal, or it is empty, or it only has a nonmicrolocal contribution. After applying a pseudodifferential cutoff described in Section 4.2 the linearized forward operator can be written as a finite sum $(\psi(y', D_{y'})F) = F_i$, where the F_i have canonical relation Λ_i. For each of the products $F_i^* F_j$ we can apply the calculus of Fourier integral operators, see for instance Hörmander [27] Theorem 25.2.3 and the preceding discussion. The order of the compose also follows from the calculus, it satisfies

$$\operatorname{order} F_i^* F_j = \operatorname{order} F_i^* + \operatorname{order} F_j + \frac{e}{2} = \frac{n-1}{2} + \frac{e}{2},$$

where e is the excess. So we have the following result.

Theorem 4.3.3 *Let C be as defined in (4.14). Suppose that* $\operatorname{rank} C$ *is maximal on $L \cap M$. Let $\psi(y', D_{y'})$ be a pseudodifferential cutoff as described in Section 4.2. Then $N = (\psi(y', D_{y'})F)^*(\psi(y', D_{y'})F)$ is the sum of a pseudodifferential operator $N_{\psi\rm do}$ of order $n-1$ and a Fourier integral operator $N_{\rm nml}$ of order $\frac{n-1}{2}$. The canonical relation $C_{\rm nml}$ of $N_{\rm nml}$ is given by (4.10), (4.11), (4.12), taking only solutions with $t \neq 0$.*

The new aspect is the simple characterisation using the matrix C. In Section 4.5 we discuss when this condition is violated and we construct an example of such a situation.

Note that the fact that the nonmicrolocal part has lower order does not imply that it is less singular as an operator between Sobolev spaces. This will be the subject of the next section.

If there are no solutions to equations (4.10) such that the rays hit the receiver resp. source manifold, then the normal operator is purely pseudodifferential. This case has been discussed by Ten Kroode e.a. [34] and by Nolan and Symes [41].

4.4 Sobolev estimates for the nonmicrolocal part and invertibility

In the previous section it is shown that under certain conditions the normal operator is the sum of a pseudodifferential operator $N_{\psi\mathrm{do}}$ of order $n-1$ and a nonmicrolocal Fourier integral operator N_{nml} of order $\frac{n-1}{2}$. The pseudodifferential part is an invertible operator $H_{(s)} \to H_{(s-n+1)}$. In this section we discuss for what values of m the nonmicrolocal part is continuous $H_{(s)} \to H_{(s-m)}$. If this is the case with $m < n-1$ then the nonmicrolocal part is less singular then the pseudodifferential part and their sum is asymptotically invertible. The infimum of the set

$$\{m \,|\, N_{\mathrm{nml}} \text{ is continuous } H_{(s)} \text{ to } H_{(s-m)}\}$$

will be called the Sobelev order of N_{nml}.

We first show that the nonmicrolocal part is at most as singular as the pseudodifferential part.

Theorem 4.4.1 *The operator F is continuous $H_{(s)} \to H_{(s-(n-1)/2)}$, and hence N is continuous $H_{(s)} \to H_{(s-n+1)}$.*

Proof We can write F as a finite sum $F = \sum_k F_k$, where the F_k have canonical relation $\Lambda_k \subset \Lambda$ such that $\Lambda_k^* \circ \Lambda_k$ is contained in the diagonal of $(T^*X\backslash 0) \times (T^*X\backslash 0)$. Then $F_k^* F_k$ is a pseudodifferential operator of order $n-1$ and hence continuous $H_{(s)} \to H_{(s-n+1)}$. Therefore the F_k are continuous $H_{(s)} \to H_{(s-(n-1)/2)}$. The "off diagonal" terms can be estimated by the diagonal terms

$$2\langle F_k u, F_l u\rangle \leq \|F_k u\|^2 + \|F_l u\|^2.$$

This implies that F is also continuous $H_{(s)} \to H_{(s-(n-1)/2)}$. □

Some basic facts about the calculation of the Sobolev order for general Fourier integral operators $\mathcal{E}'(\bar{X}) \to \mathcal{D}'(X)$ are described in Hörmander [27], Section 25.3 (here $\bar{X}$ is a copy of X such that $\bar{x} \in \bar{X}$). It is well known that if the canonical relation C is the graph of a bijective canonical map $T^*X\backslash 0 \to T^*\bar{X}\backslash 0$, then the Sobolev order equals the order of the Fourier integral operator. In that case the projections of C on $T^*\bar{X}\backslash 0, T^*X\backslash 0$ are both bijective. If A is such a Fourier integral operator of order m then A^*A is a pseudodifferential operator of order $2m$ and hence A is a continuous map $H_{(s)} \to H_{(s-m)}$. If the principal symbol of such an operator is invertible it is called *elliptic*.

If the canonical relation is not the graph of a bijective canonical map, then the projections $\bar{\pi}, \pi$ of C on $T^*\bar{X}\backslash 0$, resp. $T^*X\backslash 0$ are not bijective. Denote by $\bar{\lambda}, \lambda$ the corank of these projections at some point in C. Because $\dim \bar{X} = \dim X$ it follows from Hörmander [27] Lemma 25.3.6 that $\bar{\lambda} = \lambda$. The number λ plays an important role in estimates of the Sobolev order. In fact Theorems 25.3.8, 25.3.9 in Hörmander [27] give the estimate

$$\frac{\max_C \lambda}{6} \leq \text{Sobolev order} - \text{order} \leq \frac{\max_C \lambda}{2}. \tag{4.20}$$

If λ is constant over C (the projection is "vertical") the right hand equality is valid. If the projection is singular only in a lower dimensional subset of C then in general the Sobolev order is better. An example where the projection is singular only in a lower dimensional subset is when both projections $\bar{\pi}, \pi$ have a singularity of fold type, in this case the left hand equality is valid.

Let us now calculate λ in the case at hand. It is convenient to use new coordinates ν, θ instead of α, β, sometimes referred to as GRT coordinates. They are defined by the following map $S^{n-1} \times S^{n-1} \to S^{n-1} \times (]0, \pi[\times S^{n-2})$, see Burridge and Beylkin [13]

$$\begin{aligned} \nu &= \frac{\alpha + \beta}{\|\alpha + \beta\|} \in S^{n-1}, \\ \theta &= \left(\arccos(\alpha \cdot \beta), \frac{-\alpha + \beta}{2\sin(\arccos(\alpha \cdot \beta)/2)}\right) \in]0, \pi[\times S^{n-2}. \end{aligned} \tag{4.21}$$

Note that our definition is different from the usual definition because usually the second component of θ is denoted by ψ, and the letter θ is used only for the first component, that we will denote by θ_1. The motivation for this transformation is that ν now denotes the direction of the vector ξ, that is given in the new coordinates by

$$\xi(x, \nu, \theta, \tau) = -\frac{2\tau\cos(\theta_1/2)}{c(x)}\nu. \tag{4.22}$$

Recall the definition of the nonmicrolocal part of the canonical relation C_{nml}, see (4.10), (4.11), (4.12). We assume that $t \neq 0$ in these equations, so that we are on the nonmicrolocal part and that $\operatorname{rank} C$ is maximal so that C_{nml} is a manifold. Let $\Delta x(\bar{x}, \bar{\nu}, \bar{\theta}, t)$ be the function Δx with α, β transformed to ν, θ, let

$$V = \{(\bar{x}, \bar{\nu}, \bar{\theta}, t) \,|\, \Delta x(\bar{x}, \bar{\nu}, \bar{\theta}, t) = 0\}. \tag{4.23}$$

From the assumption that $\operatorname{rank} C$ is maximal it follows that V is a submanifold of $\bar{X} \times S^{n-1} \times (]0, \pi[\times S^{n-2}) \times \mathbb{R}$. Let $x(\bar{x}, \bar{\nu}, \bar{\theta}, t), \xi(\bar{x}, \bar{\nu}, \bar{\theta}, t, \tau)$ be the values of x, ξ for the point in C_{nml} given by $(\bar{x}, \bar{\nu}, \bar{\theta}, t, \tau)$ (see (4.11), (4.12)), and let $\bar{\xi}(\bar{x}, \bar{\nu}, \bar{\theta}, \tau)$ be given by (4.22). Then C_{nml} is obtained by mapping V into $T^*\bar{X}\backslash 0 \times T^*X\backslash 0$ as follows

$$C_{\text{nml}} = \{(x(\bar{x}, \bar{\nu}, \bar{\theta}, t), \xi(\bar{x}, \bar{\nu}, \bar{\theta}, t, \tau); \bar{x}, \bar{\xi}(\bar{x}, \bar{\nu}, \bar{\theta}, \tau) \,|\, (\bar{x}, \bar{\nu}, \bar{\theta}, t) \in V, \tau \in \mathbb{R}\}. \tag{4.24}$$

It turns out that λ equals the corank of the projection of V on $\bar{x}, \bar{\nu}$ coordinates. To see this let V be parametrized by coordinates $p = (\bar{x}', \bar{\nu}', \bar{\theta}')$, that are a subset of $(\bar{x}, \bar{\nu}, \bar{\theta}, t)$. Then C_{nml} can be parametrized by p, τ and the corank of the projection $\bar{\pi}$ is given by $\lambda = \operatorname{corank} \frac{\partial(\bar{x}, \bar{\xi})}{\partial(p,\tau)}$. Using first that the map $\bar{\xi} \to (\bar{\nu}, \|\bar{\xi}\|)$ is a diffeomorphism and then that $\frac{\partial \|\xi\|}{\partial \tau} \neq 0$, $\frac{\partial(\bar{x},\bar{\nu})}{\partial \tau} = 0$ one obtains

$$\lambda = \operatorname{corank} \frac{\partial(\bar{x}, \bar{\xi})}{\partial(p,\tau)} = \operatorname{corank} \frac{\partial(\bar{x}, \bar{\nu}, \|\xi\|)}{\partial(p,\tau)} = \operatorname{corank} \frac{\partial(\bar{x}, \bar{\nu})}{\partial p}. \tag{4.25}$$

We state the resulting estimate as a theorem.

Theorem 4.4.2 *Let the function Δx, V and the coordinates p for V be as defined above. The corank $\lambda(p)$ of the projection $\bar{\pi}$ of C_{nml} on $T^*\bar{X}\backslash 0$ at p is given by (4.25). It is also given by $\lambda = \operatorname{corank} \frac{\partial \Delta x}{\partial(\bar{\theta},t)}(p)$. The Sobolev order of N_{nml} satisfies*

$$\frac{n-1}{2} + \frac{\max_p \lambda(p)}{6} \leq \text{Sobolev order } N_{\mathrm{nml}} \leq \frac{n-1}{2} + \frac{\max_p \lambda(p)}{2}.$$

Proof We only have to prove the new formula for λ. Let $p = (\bar{x}', \bar{\nu}', \bar{\theta}')$, be coordinates for V, denote the remaining variables by $(\bar{x}'', \bar{\nu}'', \bar{\theta}'', t)$. We have

$$\frac{\partial(\bar{x}, \bar{\nu})}{\partial p} = \frac{\partial(\bar{x}', \bar{\nu}', \bar{x}'', \bar{\nu}'')}{\partial(\bar{x}', \bar{\nu}', \bar{\theta}')} = \begin{pmatrix} I & 0 \\ \frac{\partial(\bar{x}'',\nu'')}{\partial(\bar{x}',\bar{\nu}')} & \frac{\partial(\bar{x}'',\bar{\nu}'')}{\partial \bar{\theta}'} \end{pmatrix}. \tag{4.26}$$

On the other hand, by the implicit function theorem we have

$$\left(\frac{\partial \Delta x}{\partial(\bar{x}'', \bar{\nu}'', \bar{\theta}'', t)} \right)^{-1} \cdot \frac{\partial \Delta x}{\partial(\bar{\theta}', \bar{\theta}'', t)} = \begin{pmatrix} \frac{\partial(\bar{x}'',\bar{\nu}'')}{\partial \bar{\theta}'} & 0 \\ \frac{\partial(\bar{\theta}'',t)}{\partial \bar{\theta}'} & I \end{pmatrix}. \tag{4.27}$$

It follows that both $\operatorname{corank} \frac{\partial(\bar{x},\bar{\nu})}{\partial p}$ and $\operatorname{corank} \frac{\partial \Delta x}{\partial(\bar{\theta},t)}$ are equal to $\operatorname{corank} \frac{\partial(\bar{x}'',\bar{\nu}'')}{\partial(\bar{\theta}')}$. □

Note that if at some value of (x, ν) the projection is maximally degenerate, then for all values of the scattering angle traveltime injectivity is violated. So indeed in that situation it is possible that two reflectors lead to data that overlap (the wavefront sets overlap). An example is given in the next section.

It is desirable to improve the result in Theorem 4.4.2. However, the mathematical theory needed to give sharp Sobolev estimates is complicated and when this work was done there were few results, mainly for $n = 2$. Phong and Stein [44, 43] give very precise estimates for certain Fourier integral operators with analytic phase functions that also possess a certain translation symmetry. Seeger [47] obtains somewhat weaker results in the case that is relevant for us, with C^∞ phase function that does not have this translation symmetry. After this work was finished we learned about the paper [48] that gives more general results, also for $n > 2$.

We describe the result of Seeger [47] and apply it to the nonmicrolocal part of the normal operator in the case $n = 2$. Seeger gives estimates for Fourier integral operators with relation C that is the conormal bundle of a codimension one submanifold M of $\bar{X} \times X$. It is assumed that M is given by $\Phi(\bar{x}, x) = 0$. The projections of M on $X, \bar{X}$ are submersive (note that these projections are not $\pi, \bar{\pi}$). The projections $\bar{\pi}, \pi$ are singular if and only if the Monge-Ampere determinant defined by

$$I(x, \bar{x}) := \det \begin{pmatrix} \frac{\partial^2 \Phi}{\partial x \partial \bar{x}} & \left(\frac{\partial \Phi}{\partial x}\right)^t \\ \frac{\partial \Phi}{\partial \bar{x}} & 0 \end{pmatrix}.$$

vanishes. This can be checked for instance by choosing coordinates such that M is the graph of a function of $\bar{x}_1, \bar{x}_2, x_1$.

Because the projection of M on $\bar{X}, X$ is submersive the sets

$$\begin{aligned} M_{\bar{x}} &= \{x \in X \,|\, (\bar{x}, x) \in M\}, \\ M^x &= \{\bar{x} \in \bar{X} \,|\, (\bar{x}, x) \in M\} \end{aligned} \tag{4.28}$$

are smooth immersed curves. Let

$$\begin{aligned} I_{\bar{x}}(x) &= I(\bar{x}, x), & x &\in M_{\bar{x}}, \\ I^x(\bar{x}) &= I(\bar{x}, x), & \bar{x} &\in M^x. \end{aligned} \tag{4.29}$$

Now M is said to satisfy a left finite type condition of order $m, m \in \{3, 4, \dots\}$ if for every $\bar{x} \in \bar{X}$ the function $I_{\bar{x}}$ vanishes of order at most $m - 2$. Similarly M satisfies a right finite type condition of order $m, m \in \{3, 4, \dots\}$ if for every $x \in X$ the function I^x vanishes of order at most $m-2$. From Theorem 1.1 in Seeger [47] using the theory of interpolation of function spaces (see e.g. Bergh and Löfström [6]) the following for us relevant result follows

Lemma 4.4.3 *Suppose $m \geq 3$ and suppose that C satisfies both a left and a right finite type condition of order m. Then a Fourier integral operator N of order μ is continuous $H_{(s)} \to H_{(s-\mu-1/2+1/m-\epsilon)}$.*

We apply this to our problem. In our situation not the canonical relation C but the set C' obtained by setting $\xi \to -\xi$ (i.e. $C' = \{(\bar{x}, \bar{\xi}; x, -\xi) \,|\, (\bar{x}, \bar{\xi}; x, \xi) \in C\}$) can be a conormal bundle. This doesn't change the conclusions. Because C' is Lagrangian some open part of C' is the conormal bundle of such a manifold M if the rank of the projection from C' to the base space $\bar{X} \times X$ is 3 (the maximal value, since the projection vanishes along the conic direction). For the moment we assume this is the case. Violation of this condition is comparable to what happens at a caustic, this is discussed below.

Let us give the finite type condition for our case. Seeger parametrizes C' by $\bar{x}, x$ in M and a parameter for the conic direction, while we use $\bar{x}, \bar{\nu}, \bar{\theta}, t \in V$ and the parameter τ for the conic direction. Instead of $M_{\bar{x}}, M^x$ we can thus define the subsets $V_{\bar{x}}, V^x \subset V$ of V with constant value of $\bar{x}$ resp. $x(\bar{x}, \bar{\nu}, \bar{\theta}, t)$. These are also smooth immersed curves. The function $I(\bar{x}, x)$ can be replaced by the product of this function with any smooth function bounded away from zero. This is a consequence of Leibnitz rule. In this way one can derive that $I(\bar{x}, x)$ can be replaced by $\det \frac{\partial(\bar{x},\bar{\nu})}{\partial p}$ and, using equations (4.26), (4.27) also by $J(\bar{x}, \bar{\nu}, \bar{\theta}, t) = \det \frac{\partial \Delta x}{\partial(\bar{\theta},t)}$. We therefore have the following improved estimate.

Theorem 4.4.4 *The canonical relation C_{nml} satisfies a left resp. right finite type condition of order $m, m \in \{3, 4, \dots\}$ if for all $\bar{x}$ resp. x the function $J(\bar{x}, \bar{\nu}, \bar{\theta}, t)$ defined above vanishes of order at most $m - 2$ on $V_{\bar{x}}$, resp. V^x. If C_{nml} satisfies both a left and a right condition of order m then*

$$N_{\text{nml}} \text{ is continuous } H_{(s)} \to H_{(s-n/2+1/m-\epsilon)}.$$

Let us now discuss what happens when the projection of C_{nml} on $\bar{X} \times X$ is lower than 3. In that case we the result is slightly modified. According to Hörmander [27], lemma 25.3.7 one can apply a symplectic coordinate transformations on $\bar{X}$ and/or X such that C' is a conormal bundle. Such coordinate transformations correspond to application of elliptic FIO's which leave the Sobolev L^2 estimates unchanged. Becuase of the transformation we obtain different sets of curves instead of $V_{\bar{x}}, V^x$, and using these curves the theorem is still valid.

4.5 Examples

We have seen above that there are in general two problems with the normal operator. First it may not be a Fourier integral operator. Second the nonmicrolocal part may be as singular as the pseudodifferential part and then it is not clear whether the normal operator, and hence the linearized forward operator, is still invertible. In this section we give examples. In a situation with translation symmetry in a part of the medium the projections $\bar{\pi}, \pi$ of the previous section may be maximally degenerate. We discuss in how far F is not invertible. We then present numerical examples of such a situation. After that we give an example where the conditions of Theorem 4.3.3 are violated. We end the section with an example of a nonmicrolocal contribution such that the nonmicrolocal part of the normal operator is less singular than the pseudodifferential part.

In our examples we only consider a part of the subsurface that contains the two scattering points of a nonmicrolocal contribution. The rest of the medium is not of interest, it can be chosen such that the rays that are involved reach the source/receiver surfaces. In the numerical examples we give a plot of the soundspeed $c(x)$ and of some rays that lead to nonmicrolocal contributions. For the first and third example we also also give two reflectors that are mapped onto each other by nonmicrolocal contributions and we give the traveltime for both reflectors as a function of say the receiver coordinate, with source coordinate fixed (on the vertical axes bounding the region of interest, these can be mapped to the traveltimes for the real source and receiver coordinates).

Assume that in a part of the subsurface the soundspeed depends only one coordinates, say x_n, $c(x) = c(x_n)$ Suppose that this function $c(x_n)$ has a minimum somewhere. Such a configuration acts as a waveguide, rays that are shot not too far from the minimum under small angles with the plane $x_n =$ constant will be deflected back towards the minimum of c. Suppose that a certain ray, shot from $x_0 = (0, \dots, 0, h)$ with an angle α, hits the plane $x_n = h$ again at time $T = T(h, \alpha)$. It follows from equations (4.5) that the quantities $c(x)\|\xi\|$ and $\xi_1, \dots, \xi_{n-1}$ are conserved so then $\xi(T) = \xi(0)$. Hence the ray is periodic in the sence that $(x(t + kT), \xi(x + kT)) = (x(t) + k(x(T) - x(0)), \xi(t))$. The curve obtained by reflecting the ray in the vertical line is also a solution to the ray equations. Let $\beta = R \cdot \alpha$, α reflected in the vertical line. Then $x(x_0, \beta, -T) = x(x_0, \alpha, T)$. It follows that the manifold defined by (4.10) is given by

$$\bar{\beta} = R \cdot \bar{\alpha}, \qquad t = T(\bar{x}_n, \bar{\alpha}).$$

So all elements of the set V defined in (4.23) satisfy $\bar{\nu} = (0, \dots, 0, 1)$ and the projection of V on the $(\bar{x}, \bar{\nu})$ variables is singular of order $n-1$, it is maximally singular. It follows that the nonmicrolocal part has Sobolev order $n-1$. According to Hörmander [27] it is continuous $H_{(s)} \to H_{(s-(n-1))})$, it is as singular as the local part.

The fact that N_{nml} is as singular as the local part does not imply that reconstruction is impossible. This would be the case if there are $\delta c_1 \neq \delta c_2$ such that $F\delta c_1 = F\delta c_2$. This seems to be difficult to obtain globally, however it can be the case for r, s in and open subsets of Σ_r, Σ_s, as we now show.

Theorem 4.5.1 *Let the medium $c(x)$ be as described above. There is some open set in $U \subset T^*Y'$ and two contributions δc_1, δc_2, such that* $\mathrm{WF}(F \cdot \delta c_1) \cap U = \mathrm{WF}(F \cdot \delta c_2) \cap U \neq \varnothing$, *and* $\mathrm{WF}(F \cdot (\delta c_1 - \delta c_2)) \cap U = \varnothing$.

Proof Assume the source and receiver surface are located inside the medium where it is still translationally symmetric, at x_n is h. This is not a restriction. Suppose we have $s_0, p_{s,0}, r_0, p_{r,0}, \bar{x}_0, \bar{\xi}_0, x_0, \xi_0$ in the situation described above (4.10). Let δc_1 be supported on a neighborhood of $\bar{x}$, and only depend on x_n around $\bar{x}$. Let δc_2 be δc_1 translated to a neighborhood of x. By symmetry

$$F \cdot \delta c_1 = F \cdot \delta c_2 \tag{4.30}$$

close to s_0, r_0. To see this consider a reflection such that $r \leftrightarrow s$.

$$\begin{aligned}(F \cdot \delta c)(r, s, t) &= (F \cdot \delta c_{\mathrm{refl}})(r_{\mathrm{refl}}, s_{\mathrm{refl}}, t) \\ &= (F \cdot \delta c_{\mathrm{refl}})(s, r, t) \\ &= (F \cdot \delta c_{\mathrm{refl}})(r, s, t).\end{aligned}$$

Now only the microlocal properties of δc around the scattering point matter. Clearly for s, r in small neighborhoods of s_0, r_0 we have $\delta c_{1,\mathrm{refl}} = \delta c_2$ close to the scattering point x. So (4.30) follows. This proves the theorem. $\square$

Let us now give a numerical examples of the situation described above in $n = 2$ dimensions. Suppose the soundspeed is given by $c(x_1, x_2) = 1 + x_2^2$. The soundspeed $c(x)$ and some rays that lead to nonmicrolocal contributions are given in Figure 4.3a,b. Reflectors with ν vertical now lead to nonmicrolocal terms, two reflectors that are connected by a nonmicrolocal contribution are given in the Figure 4.3b. In Figure 4.3c we have drawn the arrivaltimes for a fixed source at position $(0, .5)$ and a set of receivers along a vertical line at $x_1 = 6$. Indeed the traveltime curves overlap.

We give an example where the condition $\operatorname{rank} C =$ maximal of Theorem 4.3.3 is violated. Assume again that we are in two dimension, that $c(x) = c(x_2)$ and that this function has a minimum somewhere such that a waveguide occurs. It is possible that the wavefront leaving $\bar{x}$ forms caustics. Part of this caustic may intersect the line $x_2 = h$, say for an angle $\bar{\alpha}$ and time t. At the caustic points we have

$$\frac{\partial x}{\partial \bar{\alpha}}(\bar{x}, \bar{\alpha}, t) = 0.$$

Define $\bar{\beta}$ by reflecting $\bar{\alpha}$ in the x_2 axis. Then the direction $\bar{\beta}$ gives a caustic at the same point on the x_1 axis for time coordinate $-t$. So

$$\frac{\partial x}{\partial \bar{\beta}}(\bar{x}, \bar{\beta}, -t) = 0.$$

Due to the translation symmetry we have

$$\frac{\partial x}{\partial \bar{x}}(\bar{x}, \bar{\alpha}, t) = I_2 = \frac{\partial x}{\partial \bar{x}}(\bar{x}, \bar{\beta}, -t).$$

So in this case only the last column of C is nonzero and $\operatorname{rank} C = 1$. A picture of the situation is given in Figure 4.4.

We construct an example where the projections $\pi, \bar{\pi}$ have maximal corank, see Figure 4.5. We choose $c(x)$ similar to the previous example, but we add a term that is not translation invariant. We have chosen $c(x_1, x_2) = 1 + x_2^2 + 0.25x_2^4 + 0.25x_2 \sin(x_1)$. One can determine the set V of (4.23) and except for some isolated points the projection of V on the $(\bar{x}, \bar{\nu})$ variables has maximal rank. Indeed we see that the traveltime curves are approximately tangent.

4.6 Generically the normal operator is invertible

Suppose $c(x) \in C^\infty(X)$ is a medium for which the normal operator is not an FIO or it is not invertible. We show that by perturbing the medium with a function $\delta c(x) \in C^\infty(X)$ the property that the normal operator is an invertible FIO in general is restored. Thus it is "unlikely" to encounter a medium for which the normal operator is not an invertible FIO, unless one only considers media with certain special properties (such as symmetries).

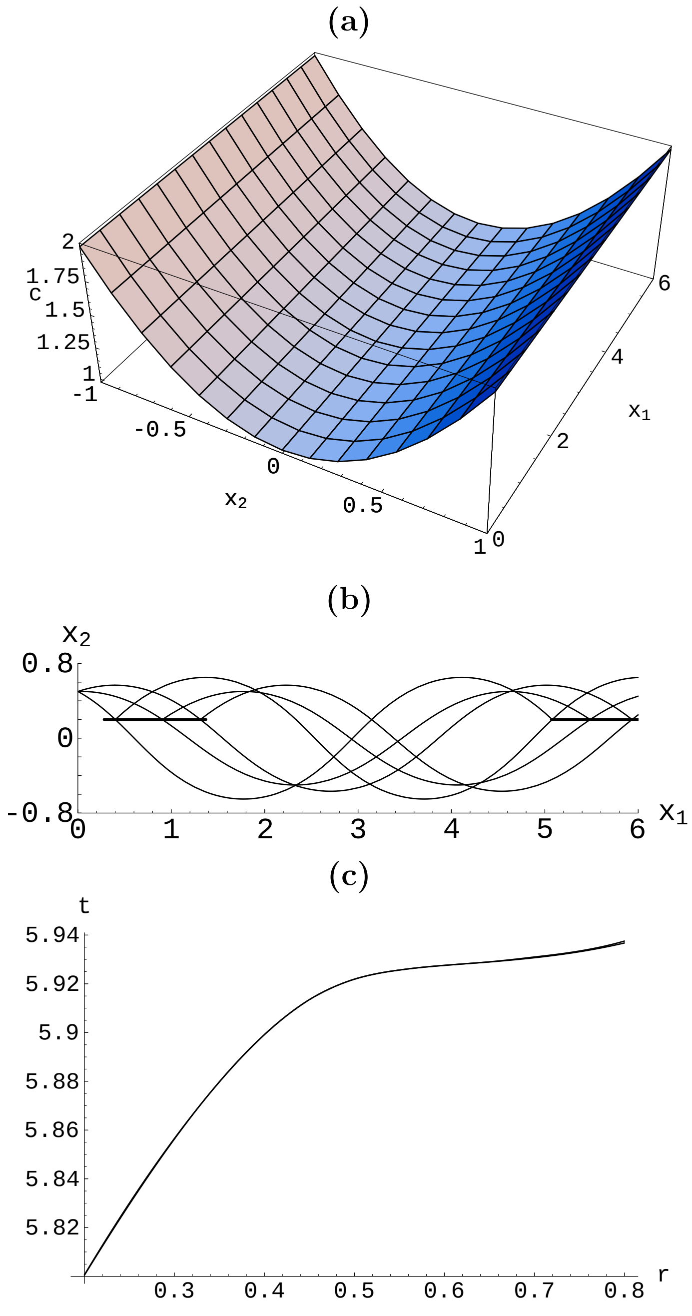


Figure 4.3: Situation with two reflectors that lead to the same arrivaltimes, i.e. the projections $\pi, \bar{\pi}$ of Section 4.4 are degenerate. (a) The soundspeed $c(x_1, x_2) = 1 + x_2^2$. (b) Some rays that lead to a nonmicrolocal contribution, two reflectors. (c) Arrival times for a source at $(0, 0.5)$ and receiver at $(6, r)$ as a function of r. The arrival times for both reflectors exactly overlap.

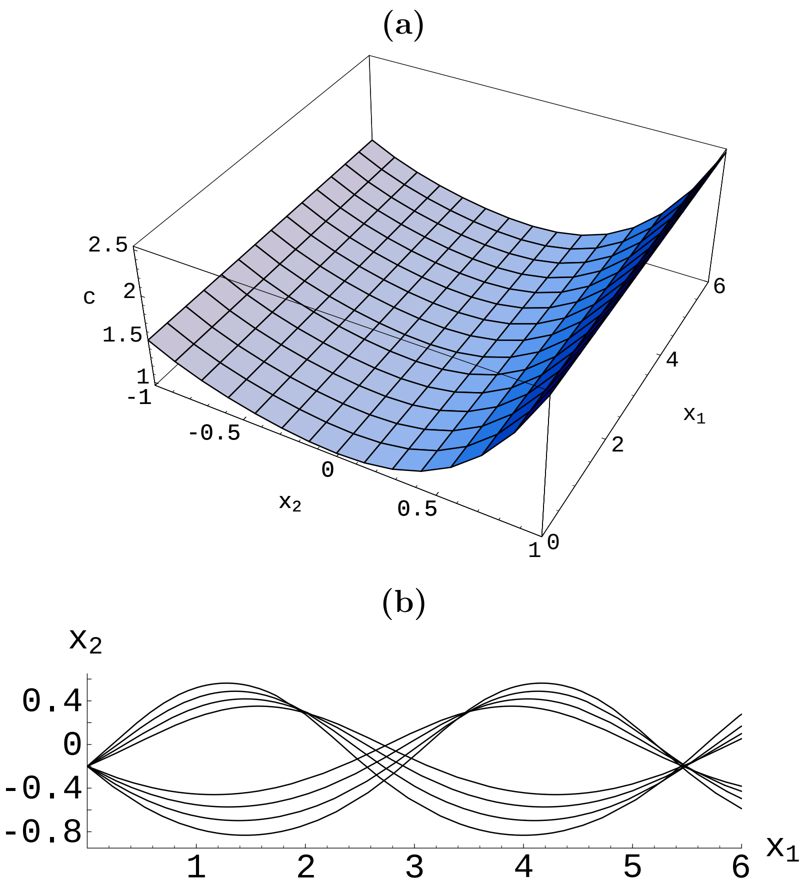


Figure 4.4: A situation where the condition of Theorem 4.3.3 is violated. (a) The soundspeed $c(x_1, x_2) = 1 + .7x_2^2 + .5x_2^3 + .3x_2^4$. (b) Ray trajectories. Intersecting caustics are formed.

Let S be the set of media $c \in C^\infty(X)$ with the property that the normal operator is a Fourier integral operator and that it is invertible. The statement is that the set S contains "almost all" of $C^\infty(X)$. Such a property is called generic. Because $C^\infty(X)$ is only a topological space and there is no measure on $C^\infty(X)$, that means that S contains a countable intersection of open dense sets (see e.g. Klingenberg [33], p. 108).

We first discuss the question whether N is a Fourier integral operator. To do this we construct a function that tests whether the criteria of Theorem 4.3.3 are satisfied. Let

$$u = (\bar{x}, \bar{\alpha}, \bar{\beta}, t, v) \in U,$$
$$U = X \times \{(\alpha, \beta) \in S^{n-1} \times S^{n-1} \,|\, \alpha + \beta \neq 0\} \times I_t \times S^{n-1}. \tag{4.31}$$

and let

$$w(u; c) = (x(\bar{x}, \bar{\alpha}, t) - x(\bar{x}, \bar{\beta}, -t), v \cdot C) \in \mathbb{R}^{4n-1}.$$

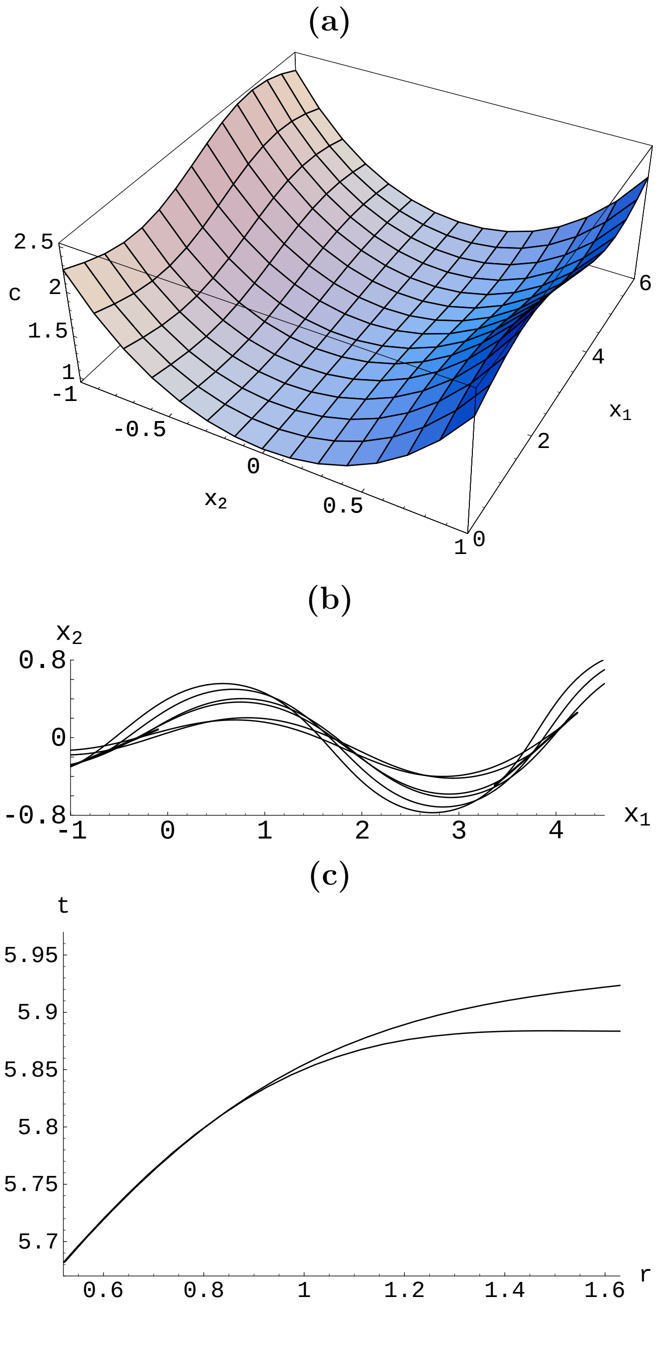


Figure 4.5: Situation with a nondegenerate nonmicrolocal contribution. (a) The soundspeed $c(x_1, x_2) = 1 + x_2^2 + 0.25x_2^4 + 0.25x_2 \sin(x_1)$. (b) Some rays that lead to a nonmicrolocal contribution, two reflectors. (c) Arrival times for a source at $(0, -0.16)$ and receiver at $(6, r)$ as a function of r. The arrival times for both reflectors are approximately tangent.

The vector v is just to test whether the rank of C is maximal. To make the dependence of w on $c = c(x)$ explicit in the notation we write $w = w(u; c)$. According to (4.12) and Theorem 4.3.3 the normal operator is an FIO if $w(u; c) \neq 0$ for all $u \in U$. Actually since we apply a microlocal cutoff to the data (see Section 4.2) the condition is that $w(u; c) \neq 0$ for all u in some compact subset $\tilde{U}$ of U.

The image of the map $u \mapsto w(u; c)$ is a subset of $\mathbb{R}^{4n-1}$ of "dimension at most 4n-2". Intuitively it is clear that such a subset in general doesn't intersect the set $\{0\}$. The important result is now Lemma 4.6.2 which shows that if intersection does occur then indeed it disappears when the coefficients are perturbed. That this is in fact the correct condition follows from the following lemma of Mather, [39], Lemma 3.2, which says so much as "transversal intersection is generic".

Lemma 4.6.1 *Let F be a topological space. Let U, W be manifolds, and V a submanifold of W. Let for each $f \in F$ there be a mapping $U \to W : u \mapsto w(u; f)$. Suppose for each $f \in F$ there exists an integer k and a continuous k-parameter family $\mathbb{R}^k \ni p \mapsto f_p$ in F, $f_{p_0} = f$, such that on a neighborhood of p_0 the map $(u, p) \mapsto w(u; f_p)$ is C^∞ and transversal to V. Then*

$$\{f \in F \,|\, u \mapsto w(u, f) \text{ is transversal to } V\}$$

is dense in F.

In the present case we take $F = C^\infty(X)$ and $V = \{0\}$. Because $\dim U + \dim V = 4n - 2 < \dim W = 4n - 1$ transversal intersection of the map $u \mapsto w(u; c)$ with V means that $w(u; c) \neq 0$ for $u \in U$. We show that in this case the assumption of Lemma 4.6.1 is satisfied.

Lemma 4.6.2 *If $w(u_0; c) = 0$ then for suitably chosen small perturbations δc_i of c the perturbations $\delta w_i(u_0; c)$ together with the $\frac{\partial w}{\partial u_j}(u_0; c)$ span $\mathbb{R}^{4n-1}$.*

Proof We first show the following. There is a finite set of variations $\delta c_1, \ldots, \delta c_n$ such that $\delta(x(\bar{x}, \bar{\alpha}, t) - x(\bar{x}, \bar{\beta}, -t))$ can be any element of $\mathbb{R}^n$. To prove this consider variations δc supported around the α ray, away from the β ray. The statement now follows from Lemma 4.A.1.

Now we show that if $v \cdot C = 0$ then there is a finite set $\delta c_1(x), \ldots, \delta c_k(x)$ of perturbations of $c(x)$, that leave the $(\bar{x}, \bar{\alpha})$ and $(\bar{x}, \bar{\beta})$ rays unchanged, while

$$v \cdot \delta C'$$

can be any element of $\mathbb{R}^{3n-2}$. Here C' is defined as the matrix C with the last column omitted. The last column $C''$of C is always nonzero, so we can choose δv such that $\delta v \cdot C''$ is nonzero. To prove that $v \cdot \delta C'$ can be any element of $\mathbb{R}^{3n-2}$ note that the inner product of v with the last column of C is $v \cdot (\alpha + \beta) = 0$. It follows that v has a nonzero component orthogonal to α, and also a nonzero component orthogonal to β. Denote these, in Fermi coordinates by $v^{\mathrm{F},\alpha}, v^{\mathrm{F},\beta}$. By Lemma 4.A.1 we can find δc

around the α and δc around the β ray such that $v^{\mathrm{F},\alpha} \cdot \delta\left(\frac{\partial(x_{\mathrm{F}},\alpha)}{\partial \bar{x}_{\mathrm{F}}}\right), v^{\mathrm{F},\beta} \cdot \delta\left(\frac{\partial(x_{\mathrm{F}},\beta)}{\partial \bar{x}_{\mathrm{F}}}\right)$ can take any values in $\mathbb{R}^{2n-2}$. We transform back to the original coordinates using the E_i, $i = 1, \ldots, n-1$. Since the E_i corresponding to the α ray together with the E_i corresponding to the β ray span $\mathbb{R}^n$ it follows that $v \cdot \delta C'$ can be any vector in $\mathbb{R}^{3n-2}$. $\square$

We can now give a mathematical proof that the property that N is a FIO is generic.

Theorem 4.6.3 *If $\tilde{U}$ is a compact subset of U (defined in (4.31)), then the set of media $c \in C^\infty(X)$ such that $w(u;c) \neq 0$ for all $u \in \tilde{U}$ is open and dense. Hence genericly the normal operator is a Fourier integral operator.*

Proof of Theorem 4.6.3 Let $u_0 \in U$. It follows from Lemma 4.6.2 that there is an open neighborhood U_0 of u_0 and parameters p such that the map $(u,p) \mapsto w(u;c_p)$ intersects $\{0\}$ transversally for $u \in U_0, p \in P$. By Lemma 4.6.1 this implies that the set S_{U_0} of $c \in C^\infty(X)$ that satisfy $w(u;c) \neq 0$ on U_0 is dense. Since the map $(c,u) \mapsto w(u;c)$ is continuous it is also open. The set $S_{\tilde{U}}$ of $c(x)$ that satisfies $w(u) \neq 0$ on $\tilde{U}$ is a finite intersection of such S_{U_0}. Therefore it is also open and dense. $\square$

Next for the case $n = 2$ we apply the same procedure to the question whether N is invertible, i.e. for the question whether the nonmicrolocal part of the normal operator has Sobolev order less than $n - 1 = 1$. We use the result of Theorem 4.4.4 and the theory described in that section. If $m \in \{3, 4, \ldots\}$ is sufficiently large then generically C_{nml} will satisfy both a left and a right finite type condition of order m. We first construct a function to test whether C_{nml} satisfies a left finite type condition of order m. Let

$$\begin{aligned} &u = (\bar{x}, \bar{\nu}, \bar{\theta}, t) \in U, \\ &U = X \times S^1 \times]0, \pi[\times]0, T_{\max}[. \end{aligned} \tag{4.32}$$

Here $\bar{\nu}, \bar{\theta}$ are coordinates as in (4.21). Let $J(\bar{x}, \bar{\nu}, \bar{\theta}, t) = \det \frac{\partial \Delta x}{\partial(\bar{\theta}, t)}(\bar{x}, \bar{\nu}, \bar{\theta}, t)$, and let v^{L} be a vectorfield on U that is parallel to $V_{\bar{x}}$ (so it depends on V and hence on the medium coefficients $c(x)$). The vectorfield v^{L} has components $(v^{\mathrm{L}}_{\bar{x}}, v^{\mathrm{L}}_{\bar{\nu}}, v^{\mathrm{L}}_{\bar{\theta}}, v^{\mathrm{L}}_{t})$, and by $(v^{\mathrm{L}})^k \cdot J$ we mean k times acting with v^{L} on J (the k-th derivative in the direction of v^{L}). Finally let

$$w(u;c) = \left(\Delta x, J, v^{\mathrm{L}} \cdot J, \ldots, (v^L)^{m-2} \cdot J\right) \in W = \mathbb{R}^{m+1} \tag{4.33}$$

According to Theorem 4.4.4 the canonical relation of N_{nml} satisfies a left finite type condition at u if $w(u;c) \neq 0$. By interchanging the roles of $\bar{x}, x$ it also satisfies a right finite type condition, and therefore if $w(u;c) \neq 0$ for all u in some compact subset $\tilde{U}$ of U then the nonmicrolocal part has Sobolev order smaller than $n - 1 = 1$ and the normal operator is invertible. We show that for m sufficiently high, in fact $m = 5$, this property is generic.

We need to prove that if $w(u;c)=0$ then its perturbation caused by some perturbation $\delta c(x)$ can take a sufficiently large set of values. This is done in the following lemma.

Lemma 4.6.4 *For suitably chosen perturbations δc_i of c the associated perturbations $\delta w_i(u_0;c)$, together with the $\frac{\partial w}{\partial u_j}(u_0;c)$ span $\mathbb{R}^{m+2}$.*

Proof Let $k \in \{0,\dots,m-1\}$. It follows from Lemma 4.A.1 that by suitable perturbations of $c(x)$, that are supported around the two rays, and vanish of order $k+1$ on the ray we can obtain an arbitrary value of the vector

$$\delta \frac{\partial^k \Delta x}{\partial \bar{\theta}^k},$$

while perturbations of lower derivatives vanish.

In the case $k=0$, we can have arbitrary $\Delta x(\bar{x},\bar{\nu},\bar{\theta},t)$. For $k=1$ we see that $J=\frac{\partial \Delta x_1}{\partial \bar{\theta}}\frac{\partial \Delta x_2}{\partial t}-\frac{\partial \Delta x_2}{\partial \bar{\theta}}\frac{\partial \Delta x_1}{\partial t}$ can have a nonzero perturbation. We use the case $k>1$ to perturb $(v^{\mathrm{L}})^{k-1}\cdot J$. One can check from the theory of Section 4.4 that when $J=0$ then $v^{\mathrm{L}}_{\bar{x}}=v^{\mathrm{L}}_{\bar{\nu}}=0$, and $(v^{\mathrm{L}}_{\bar{\theta}},v^{\mathrm{L}}_t)\in \ker \frac{\partial \Delta x}{\partial(\bar{\theta},t)}$. To $\delta(v^{\mathrm{L}})^{k-1}J$ only the term involving the highest order derivatives contributes, it is given by

$$\delta \sum_{i_1,\dots,i_{k-1}} (v_{\bar{\theta}},v_t)_{i_1}\dots(v_{\bar{\theta}},v_t)_{i_{k-1}}\frac{\partial}{\partial(\bar{\theta},t)_{i_1}}\cdots\frac{\partial}{\partial(\bar{\theta},t)_{i_1}}\left(\frac{\partial \Delta x_1}{\partial \bar{\theta}}\frac{\partial \Delta x_2}{\partial t}-\frac{\partial \Delta x_2}{\partial \bar{\theta}}\frac{\partial \Delta x_1}{\partial t}\right).$$

This term can be made nonzero. So by successively choosing perturbations vanishing of order $k=0,\dots,m-1$ we can have all components of δw arbitrary. □

Now for $m=5$ the image of the map $u\mapsto w(u;c)$ is of lower dimension then the space $W=\mathbb{R}^{m+1}$. The following theorem states the N is generically invertible. The proof is omitted, because it parallels that of Theorem 4.6.3 with Lemma 4.6.2 replaced by Lemma 4.6.4

Theorem 4.6.5 *Let $m=5$. If $\tilde{U}$ is some compact subset of U (defined in (4.32)) then the set of media $c\in C^\infty(X)$ such that $w(u;c)\neq 0$ for $u\in\tilde{U}$ is open and dense. This implies that generically the normal operator is asymptotically invertible.*

4.7 Discussion

Let us discuss how we have improved the understanding of the nonmicrolocal terms to the normal operator.

First recall from the introduction and Section 4.5 that in the case of nonmicrolocal contributions there can be scatterers at different positions in the subsurface such that the diffractorsurfaces (the singular support of the data) are tangent, see e.g. Figures 4.3, 4.5. By imaging straightforwardly the data corresponding to one scatterer may be imaged to the position of the other scatterer. Often this artifact will be smaller then the correct image, because of “stacking”, i.e. an integration over the redundant direction, that suppresses events that do not “line up”. The results of Section 4.4 in fact quantify this effect of stacking, at least the highfrequency behavior. If the diffractorsurfaces overlap then the artifact can have same order of magnitude as the correct image.

There were two main steps. First we proved under which conditions the nonmicrolocal part of the normal operator N_{nml} is an FIO. This allows us to analyze it. It is essentially determined by the function $\Delta x(\bar{x}, \bar{\alpha}, \bar{\beta}, t)$ defined in (4.10), and we did not have to consider different type of caustics as in Ten Kroode e.a. [34]. Secondly we analyzed the Sobolev regularity of N_{nml}. The Sobolev regularity is essentially determined by the set V defined in (4.23) and the projection of V on the variables $\bar{x}, \bar{\nu}$ (in other words by the θ dependence of the nonmicrolocal terms).

The examples give more insight in when nonmicrolocal terms can occur and in how far they are relevant. It seems that in previous work people haven’t realized that in the example of Figure 4.3 the nonmicrolocal terms (artifacts) are in fact of the same size as the pseudodifferential part (the correct image).

It would be interesting to see whether in partial acquisition geometries (see Nolan and Symes [41]) such degerate examples occur more often. In this case the imaging could be affected more severely, because there is less redundancy than in our case.

4.A Perturbations of a ray due to medium perturbations

In Section 4.6 it is necessary to know how the rays and its derivatives are changed when the medium is perturbed. This will be discussed in this appendix. This material is not new, but we couldn’t find a reference that describes such results for acoustic media. A discussion for the case where the soundspeed $c(x)$ is replaced by a general Riemannian metric can be found in Klingenberg [33], Section 3.3. We use some Riemannian geometry, so in this section we will use upper indices for the coordinates, and lower indices for the cotangent vectors.

The rays are the geodesics corresponding to the metric $g^{ij} = c(x)^2\delta^{ij}$. This is a conformal metric (proportional to δ^{ij}). The rays can be calculated by solving the Hamilton system associated to the Hamiltonian $H(x, \xi) = \sqrt{\sum_{i,j} g^{ij}\xi_i\xi_j} = c(x)\|\xi\|$, see (4.4). The square root is taken in the definition of H so that the parameter t is the traveltime (the arclength). The solution with initial values x_0, ξ_0 and parameter t will be denoted by $(x(x_0, \xi_0, t), \xi(x_0, \xi_0, t))$.

The derivative $\frac{\partial^m(x,\xi)}{\partial(x_0,\xi_0)^m}(x_0, \xi_0, t)$ can be calculated by solving the ordinary differential system

$$\frac{\partial}{\partial t}\frac{\partial^m(x,\xi)}{\partial(x_0,\xi_0)^m}(x_0,\xi_0,t) = \frac{\partial^m}{\partial(x_0,\xi_0)^m}\left(\frac{\partial H}{\partial \xi}(x(x_0,\xi_0,t),\xi(x_0,\xi_0,t)),\right.$$
$$\left. - \frac{\partial H}{\partial x}(x(x_0,\xi_0,t),\xi(x_0,\xi_0,t))\right). \tag{4.34}$$

By applying the derivatives we see that this is a linear system with homogeneous term similar to the ordinary Jacobi equation

$$\begin{pmatrix} \frac{\partial^2 H}{\partial \xi \partial x} & \frac{\partial^2 H}{\partial \xi^2} \\ -\frac{\partial^2 H}{\partial x^2} & -\frac{\partial^2 H}{\partial x \partial \xi} \end{pmatrix} \cdot \frac{\partial^m(x,\xi)}{\partial(x_0,\xi_0)^m}.$$

The inhomogeneous term consists of products of lower derivatives $\frac{\partial^k(x,\xi)}{\partial(x_0,\xi_0)^k}$ with higher derivatives of H. The Jacobi equations (4.5) hence play a special role, and will discussed first, before we look at the the effect of perturbations of c.

The Jacobi equations can be simplified considerably by using coordinates that are centred around the ray, so called Fermi coordinates[1]. Fermi coordinates are described for instance in Klingenberg [33]. The new coordinates consist of a coordinate s that denotes the time along the ray and coordinates $x_{\mathrm{F}}^1, \ldots, x_{\mathrm{F}}^{n-1}$ that denote the distance from the ray, in units of time. To define them let $E_0, \ldots, E_{n-1}$ be a set of orthonormal vectors (with respect to g^{ij}) in $T_{x_0}X$, such that $E_0 = \frac{\partial x}{\partial t}(x_0,\xi_0,0)$. Denote by $E_i(t)$ the E_i parallel transported along the ray. Consider now the map

$$\Psi : (s, x_{\mathrm{F}}^1, \ldots, x_{\mathrm{F}}^{n-1}) \mapsto \exp_{x(x_0,\xi_0,s)}\left(\sum_{i=1}^{n-1} E_i(s) x_{\mathrm{F}}^i\right).$$

When the x_{F}^i are sufficiently close to 0 this map defines a set of coordinates around the ray. The transformation matrix can be written

$$\frac{\partial x^i}{\partial s}(s,0) = E_0^i(s), \qquad \frac{\partial x^i}{\partial x_{\mathrm{F}}^j}(s,0) = E_j^i(s),$$

where E_j^i denotes the i-th component of E_j. The corresponding cotangent vectors will be denoted by $\sigma, \xi_i^{\mathrm{F}}$. Note that if ξ is close to ξ_0 then σ is approximately equal to the length of ξ, while while $\frac{\xi_i^{\mathrm{F}}}{\sigma}$ parametrizes the angle with ξ_0. From the definition it follows that for the new metric g_{ij}^{F}

$$g_{ij}^{\mathrm{F}}(s,0) = \delta_{ij}, \qquad \frac{\partial g_{ij}^{\mathrm{F}}}{\partial x_{\mathrm{F}}^k}(s,0) = 0,$$
$$\frac{\partial^2 g_{00}^{\mathrm{F}}}{\partial x_{\mathrm{F}}^i x_{\mathrm{F}}^j}(s,0) = -2R_{0i0j}^{\mathrm{F}}(s,0), \tag{4.35}$$

[1] In the geophysical literature one can find so called ray centred coordinates that have similar properties, see e.g. Popov and Pšenčík [45].

where the index 0 corresponds to the s coordinate and R_{ijkl} is the Riemann curvature tensor.

To obtain the ray and its derivatives in the new coordinates we can just set $H(s, x_{\mathrm{F}}, \sigma, \xi^{\mathrm{F}}) = \sqrt{\sum_{i=0}^{n-1} g_{\mathrm{F}}^{ij} \xi_i^{\mathrm{F}} \xi_j^{\mathrm{F}}}$. Obviously the ray is given by $s(t) = t, \sigma(t) = \sigma_0, x_{\mathrm{F}}(t) = 0, \xi^{\mathrm{F}}(t) = 0$. We set $\sigma_0 = 1$. When we now do the calculation for the Jacobi equations we find that the nonzero parts of the Jacobi matrix are $\frac{\partial s}{\partial s_0} = 1$, $\frac{\partial \sigma}{\partial \sigma_0} = 1$ and $\frac{\partial(x_{\mathrm{F}}, \xi^{\mathrm{F}})}{\partial(x_{\mathrm{F},0}, \xi^{\mathrm{F},0})}$ that satisfies the following ODE along the ray

$$\frac{\partial}{\partial t} \frac{\partial(x_{\mathrm{F}}, \xi^{\mathrm{F}})}{\partial(x_{\mathrm{F},0}, \xi^{\mathrm{F},0})}(t) = \begin{pmatrix} 0 & I_{n-1} \\ A(t) & 0 \end{pmatrix} \frac{\partial(x_{\mathrm{F}}, \xi^{\mathrm{F}})}{\partial(x_{\mathrm{F},0}, \xi^{\mathrm{F},0})}(t), \tag{4.36}$$

where

$$A_{ij}(s) = \tfrac{1}{2} \frac{\partial^2 g_{00}^{\mathrm{F}}}{\partial x_{\mathrm{F}}^i \partial x_{\mathrm{F}}^j} = -R_{i0j0}^{\mathrm{F}}(s, 0).$$

To obtain an expression for the matrix A_{ij} in terms of the original quantity $c(x)$ one first computes the Riemann curvature in the original coordinates. Then the transformation rules are used to obtain the curvature in the new coordinates. This leads to the following result

$$A_{ij} = -\delta_{ij} c^{-1} \frac{\partial^2 c}{\partial s^2} - c^{-1} \frac{\partial^2 c}{\partial x_{\mathrm{F}}^i \partial x_{\mathrm{F}}^j} + \text{lower derivatives of } c.$$

So we now have a precise description of the matrix $\frac{\partial(s, x_{\mathrm{F}})}{\partial(s_0, x_{\mathrm{F},0})}$.

Now assume that the soundspeed is perturbed by some smooth function that we denote by $\delta c(x)$ (this is not the same as the highfrequency $\delta c(x)$ of the Born approximation). We investigate the corresponding first order perturbations $\delta s(t)$, $\delta x_{\mathrm{F}}(t)$, $\delta \xi^{\mathrm{F}}(t)$. To simplify the notation we define $z_\kappa = (x_{\mathrm{F}}^1, \ldots, x_{\mathrm{F}}^{n-1}, \xi_1^{\mathrm{F}}, \ldots, \xi_{n-1}^{\mathrm{F}})$, $J_{\kappa\lambda} = \begin{pmatrix} 0 & I_{n-1} \\ -I_{n-1} & 0 \end{pmatrix}$.

Lemma 4.A.1 *A perturbation deltapc(x) of the medium that is nonzero on the ray results in a nonzero perturbation $\delta s(t)$. One can obtain arbitrary $\delta s(t)$ by choosing δc suitably. A perturbation that vanishes of order $m+1$ on the ray ($m = 0, 1, \ldots$) results in a perturbation $\delta \frac{\partial^m z}{\partial z_0^m}$, while $\delta s(t) = 0$ and $\delta \frac{\partial^k z}{\partial z_0^k} = 0$ for $k < m$. Let $v_1, \ldots, v_m \in \mathbb{R}^{2n-2}$ and let $(v_1, \ldots, v_m) \cdot \delta \frac{\partial^m z}{\partial z_0^m} \in \mathbb{R}^{2n-2}$ be some contraction of $\delta \frac{\partial^m z}{\partial z_0^m}$ with the vectors $v_1, \ldots, v_m$. One can obtain arbitrary values of $(v_1, \ldots, v_m) \cdot \delta \frac{\partial^m z}{\partial z_0^m}(t) \in \mathbb{R}^{2n-2}$ by choosing $\delta c(x)$ suitably.*

Proof If we have a perturbation $\delta c(x)$, then

$$\delta g^{ij} = 2c\, \delta c\, \delta^{ij}, \qquad \delta g_{\mathrm{F}}^{ij} = \sum_{k,l} \frac{\partial x_{\mathrm{F}}^i}{\partial x^k} \frac{\partial x_{\mathrm{F}}^j}{\partial x^l} 2c\, \delta c\, \delta^{kl},$$

and

$$\delta H = \frac{1}{2\sqrt{\sum_{i,j} g_{\mathrm{F}}^{ij} \xi_i^{\mathrm{F}} \xi_j^{\mathrm{F}}}} \sum_{k,l} \xi_k^{\mathrm{F}} \xi_l^{\mathrm{F}} \, \delta g_{\mathrm{F}}^{kl}.$$

Now suppose we set

$$\delta c(s, x_{\mathrm{F}}) = c(s, x_{\mathrm{F}}) \, \delta A(s),$$

close to the ray, going smoothly to zero away from the ray. In that case we find $\frac{\partial}{\partial t}\delta s(t) = \frac{\delta c(s)}{c(s)}$, and hence

$$\delta s(t) = \int_0^t \delta A(t') \, \mathrm{d}t'.$$

So by this choice of δc it is possible to obtain an arbitrary perturbation in the direction along the ray.

Next suppose δc is of the form

$$\delta c(s, x_{\mathrm{F}}) = c(s, x_{\mathrm{F}}) \sum_{k_1, \dots, k_{m+1}} \delta A_{k_1 \dots k_{m+1}}(s) x_{\mathrm{F}}^{k_1} \dots x_{\mathrm{F}}^{k_{m+1}}, \tag{4.37}$$

close to the ray, again going smoothly to zero away from the ray. For the perturbation $\delta \frac{\partial^m z}{\partial {z_0}^m}$ we find

$$\begin{aligned} \frac{\partial}{\partial t} \delta \frac{\partial^m z_\kappa}{\partial z_{0,\lambda_1} \dots \partial z_{0,\lambda_m}} = & \sum_{\mu, \nu_1, \dots, \nu_m} J_{\kappa\mu} \frac{\partial^{m+1} \delta H}{\partial z_\mu \partial z_{\nu_1} \dots \partial z_{\nu_m}} \frac{\partial z_{\nu_1}}{\partial z_{0,\lambda_1}} \dots \frac{\partial z_{\nu_m}}{\partial z_{0,\lambda_m}} + \\ & + \sum_{\mu,\nu} J_{\kappa\mu} \frac{\partial^2 H}{\partial z_\mu \partial z_\nu} \, \delta \frac{\partial^m z_\nu}{\partial z_{0,\lambda_1} \dots \partial z_{0,\lambda_m}}. \end{aligned} \tag{4.38}$$

There are no other nonzero contributions since $\delta \frac{\partial^k z}{\partial {z_0}^k}$ vanishes for $k < m$ and $\frac{\partial^k \delta H}{\partial z^k}$ vanishes for $k < m + 1$. The only nonzero part of $\frac{\partial^{m+1} \delta H}{\partial z^{m+1}}$ is when all the derivatives are with respect to x_{F} and they all act on δc, so

$$\frac{\partial^{m+1} \delta H}{\partial z_{\mu_1} \dots \partial z_{\mu_{m+1}}} = \delta A_{\mu_1 \dots \mu_{m+1}},$$

where we define $\delta A_{\mu_1 \dots \mu_{m+1}}$ to be 0 if any of the indices $\mu_i > n - 1$. Let $\Phi_{\kappa\lambda}(t, t')$ be the fundamental solution to (4.36). The solution of the differential equation is

$$\delta \frac{\partial^m z_\kappa}{\partial z_{0,\lambda_1} \dots \partial z_{0,\lambda_m}}(t) = \int_0^t \sum_{\mu, \nu_1 \dots \nu_m, \rho} \Phi_{\kappa\mu}(t, t') J_{\mu\rho} \delta A_{\rho\nu_1 \dots \nu_m} \prod_{i=1}^m \Phi_{\nu_i \lambda_i}(t', 0) \, \mathrm{d}t'. \tag{4.39}$$

By choosing $\delta A_{k_1\ldots k_{m+1}}$ in different ways a large set of values $\delta\frac{\partial^m z}{\partial {z_0}^m}$ can be obtained. To see this rewrite the solution as

$$\frac{\partial^m z_\kappa}{\partial z_{0,\lambda_1}\ldots\partial z_{0,\lambda_m}}(t) = \sum_{\sigma_1,\nu_1,\ldots,\nu_m} \delta Z_{\kappa\nu_1\ldots\nu_m} \prod_{i=1}^m \Phi_{\nu_i\lambda_i}(t,0),$$
$$\delta Z_{\kappa\lambda_1\ldots\lambda_m} = \int_0^t \sum_{\mu,\rho,\nu_1,\ldots,\nu_m} \Phi_{\kappa\mu}(t,t')J_{\mu\rho}\delta A_{\rho\nu_1\ldots\nu_m}(t') \prod_{i=1}^m \Phi_{\nu_i\lambda_i}(t',0)\,\mathrm{d}t'. \tag{4.40}$$

If $t-t'$ is small then

$$\Phi(t,t') = \begin{pmatrix} I_{n-1} & (t-t')I_{n-1} \\ (t-t')A(t) & I_{n-1} \end{pmatrix}.$$

up to higher order terms (by this we mean that for each subblock there may be higher order terms of different order). Now suppose that δA is supported in a small interval (t_1,t). Then, to highest order, we can give an explicit expression for δZ. In the following we write $\kappa = k+\alpha(n-1), \lambda_i = l_i + \beta_i(n-1),\ \alpha,\beta_i \in \{0,1\}$, $k, l_i \in \{1,\ldots,n-1\}$ to indicate whether κ, λ_i refer to x or to ξ coordinates. With this notation

$$\delta Z_{\kappa\lambda_1\ldots\lambda_m} = -\int_{t_1}^t (-1)^{1+\beta_1+\ldots+\beta_m}(t-t')^{1-\alpha+\beta_1+\ldots+\beta_m}\delta A_{kl_1\ldots l_m}(t')\,\mathrm{d}t'.$$

The tensor δZ is not the general tensor, but it follows from this equation that at least the contraction of δZ with m vectors $\in \mathbb{R}^{2n-2}$ is the general element of $\mathbb{R}^{2n-2}$. It follows that the same is true for $\delta\frac{\partial^m z}{\partial z_0^m}(t)$. This proves the lemma. $\square$

Bibliography

[1] R. A. Adams. *Sobolev spaces.* Academic Press, New York, 1975.

[2] K. Aki and P. G. Richards. *Quantitative Seismology: Theory and Methods*, volume 1. Freeman, San Francisco, 1980.

[3] K. Aki and P. G. Richards. *Quantitative Seismology: Theory and Methods*, volume 2. Freeman, San Francisco, 1980.

[4] G. Bao. Smoothness between coefficients and boundary values for the wave equation. *SIAM J. Math. Anal.*, 29(2):380–394, 1998.

[5] G. Bao and W. W. Symes. On the sensitivity of solutions of hyperbolic equations to the coefficients. *Comm. Partial Differential Equations*, 21:395–422, 1996.

[6] J. Bergh and J. Löfström. *Interpolation Spaces, An Introduction.* Springer-Verlag, Berlin, 1976.

[7] G. Beylkin. The inversion problem and applications of the generalized Radon transform. *Comm. Pure Appl. Math.*, 37:579–599, 1984.

[8] G. Beylkin. Imaging of discontinuities in the inverse scattering problem by inversion of a causal generalized Radon transform. *J. Math. Phys.*, 26:99–108, 1985.

[9] G. Beylkin and R. Burridge. Linearized inverse scattering problems in acoustics and elasticity. *Wave Motion*, 12:15–52, 1990.

[10] N. Bleistein. On the imaging of reflectors in the earth. *Geophysics*, 52:931–942, 1987.

[11] N. Bourbaki. *Intégration*, volume VI of *Éléments de mathématique.* Hermann, Paris, second edition, 1965.

[12] P. J. Braam and J. J. Duistermaat. Normal forms of real symmetric systems with multiplicity. *Indag. Mathem.*, 4:197–232, 1993.

[13] R. Burridge and G. Beylkin. On double integrals over spheres. *Inverse Problems*, 4(1):1–10, 1988.

[14] C. H. Chapman. Reflection/transmission coefficient reciprocities in anisotropic media. *Geophys. J. Internat.*, 116:498–501, 1994.

[15] F. Colombini and S. Spagnolo. On the convergence of solutions of hyperbolic equations. *Comm. Partial Differential Equations*, 3:77–103, 1978.

[16] N. Dencker. On the propagation of polarization sets for systems of real principal type. *J. Funct. Anal.*, 46:351–372, 1982.

[17] J. J. Duistermaat. *Fourier Integral Operators.* Birkhäuser, Boston, 1996.

[18] N. Dunford and J. T. Schwartz. *Linear Operators*, volume 1, General Theory. Interscience, New York, 1958.

[19] D. M. Elton. *Hyperbolic Partial Differential Equations with Singular Coefficients.* PhD thesis, University of Oxford, 1997.

[20] H. O. Fattorini. *Second Order Linear Differential Equations in Banach Spaces.* North-Holland, Amsterdam, 1985.

[21] N. L. Frazer and M. K. Sen. Kirchhoff-Helmholtz reflection seismograms in a laterally inhomogeneous multi-layered medium – I. Theory. *Geophys. J. R. Astr. Soc.*, 80:121–147, 1985.

[22] S. Hansen. Solution of a hyperbolic inverse problem by linearization. *Comm. Partial Differential Equations*, 16:291–309, 1991.

[23] M. V. de Hoop and N. Bleistein. Generalized Radon transform inversions for reflectivity in anisotropic elastic media. *Inverse Problems*, 13:669–690, 1997.

[24] M. V. de Hoop and S. Brandsberg-Dahl. Maslov asymptotic extension of generalized Radon transform inversion in anisotropic elastic media: a least-squares approach. *Inverse Problems*, 16:519–562, 2000.

[25] L. Hörmander. *The Analysis of Linear Partial Differential Operators*, volume 1. Springer-Verlag, Berlin, 1983.

[26] L. Hörmander. *The Analysis of Linear Partial Differential Operators*, volume 3. Springer-Verlag, Berlin, 1985.

[27] L. Hörmander. *The Analysis of Linear Partial Differential Operators*, volume 4. Springer-Verlag, Berlin, 1985.

[28] A. E. Hurd and D. H. Sattinger. Questions of existence and uniqueness for hyperbolic equations with discontinuous coefficients. *Trans. Amer. Math. Soc.*, 132:159–174, 1968.

[29] V. Isakov. *Inverse Problems for Partial Differential Equations.* Springer-Verlag, New York, 1998.

[30] V. Y. Ivrii. Wave fronts of solutions of symmetric pseudodifferential systems. *Siberian Math. J.*, 20:390–405, 1979.

[31] B. L. N. Kennett. *Seismic Wave Propagation in Stratified Media.* Cambridge University Press, Cambridge, 1983.

[32] A. Kirsch. *An Introduction to the Mathematical Theory of Inverse Problems.* Springer-Verlag, New York, 1996.

[33] W. Klingenberg. *Lectures on Closed Geodesics.* Springer-Verlag, Berlin, 1978.

[34] A. P. E. ten Kroode, D. J. Smit, and A. R. Verdel. A microlocal analysis of migration. *Wave Motion*, 28:149–172, 1998.

[35] L. D. Landau and E. M. Lifshitz. *Theory of elasticity.* Pergamon Press, London, 1959.

[36] J. L. Lions and E. Magenes. *Non-Homogeneous Boundary Value Problems and Applications*, volume 1. Springer-Verlag, Berlin, 1972.

[37] J. E. Marsden and T. J. R. Hughes. *Mathematical Foundations of Elasticity.* Prentice-Hall, Englewood Cliffs (NJ), 1983.

[38] V. P. Maslov and M. V. Fedoriuk. *Semi-Classical Approximation in Quantum Mechanics.* D. Reidel Publishing Company, 1981.

[39] J. N. Mather. Stability of C^∞ mappings: V, transversality. *Advances in Math.*, 4:301–336, 1970.

[40] M. J. P. Musgrave. *Crystal Acoustics.* Holden-Day, 1970.

[41] C. J. Nolan and W. W. Symes. Global solution of a linearized inverse problem for the wave equation. *Comm. Partial Differential Equations*, 22(5-6):919–952, 1997.

[42] A. Pazy. *Semigroups of Linear Operators and Applications to Partial Differential Equations.* Springer-Verlag, New York, 1983.

[43] D. H. Phong and E. M. Stein. Radon transforms and torsion. *Int. Math. Res. Notices*, 4:49–60, 1991.

[44] D. H. Phong and E. M. Stein. The Newton polyhedron and oscillatory integral operators. *Acta Math.*, 179:105–152, 1997.

[45] M. M. Popov and Pšenčík. Ray amplitudes in inhomogeneous media with curved interfaces. *Geofysikální Sborník*, 24:111–129, 1976.

[46] Rakesh. A linearized inverse problem for the wave equation. *Comm. Partial Differential Equations*, 13(5):573–601, 1988.

[47] A. Seeger. Degenerate Fourier integral operators in the plane. *Duke Math. J.*, 71:685–745, 1993.

[48] A. Seeger. Radon transforms and finite type conditions. *J. Amer. Math. Soc.*, 11(4):869–897, 1998.

[49] R. E. Sheriff and L. P. Geldart. *Exploration Seismology.* Cambridge University Press, second edition, 1995.

[50] S. Spagnolo. Sul limite delle soluzioni di problemi di Cauchy relativi all'equazione del calore. *Ann. Scuola Norm. Sup. Pisa*, 21:657–699, 1967.

[51] S. Spagnolo. Sulla convergenza di soluzioni di equazioni paraboliche ed ellittiche. *Ann. Scuola Norm. Sup. Pisa*, 22:571–597, 1968.

[52] C. C. Stolk. Microlocal analysis of a seismic linearized inverse problem. *Wave Motion*, 32:267–290, 2000.

[53] C. C. Stolk and M. V. de Hoop. Modeling and inversion of seismic data in anisotropic elastic media. Preprint 1143, Department of Mathematics, Utrecht University, March 2000.

[54] W. A. Strauss. On continuity of functions with values in various banach spaces. *Pacific J. Math.*, 19(3):543–551, 1966.

[55] W. W. Symes. A differential semblance algorithm for the inverse problem of reflection seismology. *Comput. Math. Appl.*, 22(4-5):147–178, 1991.

[56] M. E. Taylor. Reflection of singularities of solutions to systems of differential equations. *Comm. Pure Appl. Math.*, 28:457–478, 1975.

[57] F. Treves. *Introduction to Pseudodifferential and Fourier Integral Operators*, volume 1. Plenum Press, New York, 1980.

[58] F. Treves. *Introduction to Pseudodifferential and Fourier Integral Operators*, volume 2. Plenum Press, New York, 1980.

[59] S. Xu, H. Chauris, G. Lambaré, and M. Noble. Common angle migration: a strategy for imaging complex media. Preprint École des Mines de Paris, 1999.

[60] O. Yilmaz. *Seismic data processing.* Society of Exploration Geophysicists, 1987.

[61] K. Yosida. *Functional Analysis.* Springer-Verlag, Berlin, 1965.

Samenvatting

In een seismisch experiment probeert men door metingen aan het aardoppervlak een beeld van de ondergrond te verkrijgen. Hiertoe worden met bronnen aan het oppervlak elastische golven in de ondergrond opgewekt. Als bronnen kunnen een explosie of een zwaar trillend voorwerp (een vibroseis truck) worden gebruikt. Vervolgens worden de golven die terugkeren naar het oppervlak geregistreerd. Het probleem is nu om uit de verkregen data een beeld van de ondergrond te reconstrueren. In dit proefschrift onderzoeken we enkele wiskundige problemen die hieruit voortkomen.

Om de data te modelleren beschouwen we de aarde als een akoestisch of als een elastisch medium. Zo'n medium wordt beschreven door mediumparameters, die in het algemeen van de positie afhangen. In het akoestische geval zijn dat de lokale geluidssnelheid en de dichtheid, in het elastische geval de elasticiteitstensor en de dichtheid. De toestand van het medium wordt beschreven door de akoestische of elastische golfvergelijking, waarin de parameters als coëfficiënten voor komen. De meetgegevens kunnen nu ruwweg als volgt worden verklaard. Vanuit de bron planten de golven zich voort de ondergrond in. Op posities waar de coëfficiënten van het medium sterk variëren wordt een gedeelte van de energie gereflecteerd, terwijl een ander deel doorgaat. Wanneer de gereflecteerde signalen weer bij het oppervlak komen worden ze geregistreerd.

We nemen aan dat de mediumcoëfficiënten discontinu kunnen variëren, bijvoorbeeld bij een overgang tussen twee lagen. De eerste vraag die we onszelf stellen is of de akoestische en elastische golfvergelijkingen oplossingen hebben in dat geval. Het blijkt uit bestaande literatuur dat dit inderdaad het geval is, dat de oplossingen uniek zijn en dat ze stabiel van de bronfunctie en de beginwaarden afhangen. De oplossingen hangen bovendien continu af van de coëfficiënten. Dit wordt beschreven in hoofdstuk 2. Daarnaast geven we enkele nieuwe resultaten voor het geval waar de coëfficiënten glad zijn, behalve langs een glad oppervlak waar ze een sprongdiscontinuïteit hebben. We onderzoeken ook de afgeleide van de oplossingen naar de coëfficiënten.

De vraag is nu of uit de data de in het algemeen discontinue coëfficiënten bepaald kunnen worden. In hoofdstukken 3 en 4 gebruiken we hoogfrequente asymptotiek, in het bijzonder de theorie van Fourier-integraaloperatoren, om dit probleem aan te pakken. In de hoogfrequente limiet, en als het medium voldoende glad is, planten de golven zich voort langs stralen. Hierop gebaseerde methoden worden al succesvol toegepast. We geven nieuwe resultaten voor elastische media en voor het geval dat de golffronten caustieken (brandfiguren) vormen. Dat wil zeggen dat de stralen die vanuit een punt in verschillende richtingen worden weggeschoten elkaar snijden.

In hoofdstuk 3 beschouwen we elastische media. We modelleren de data zowel met de Born- als met de Kirchhoff-benadering. In de Born-benadering worden de mediumparameters geschreven als de som van een glad achtergrond medium, en een singuliere (hoogfrequente) verstoring. De reflecties worden dan beschreven door de corresponderende verstoring van de Greense functie. In de Kirchhoff-benadering neemt men aan dat het medium stuksgewijs glad is, met sprongen langs gladde oppervlakken. Aan de sprongoppervlakken wordt een gedeelte van de inkomende golven gereflecteerd. Onder bepaalde aannames kunnen we in deze gevallen het singuliere deel van de mediumcoëfficiënten reconstrueren. We geven ook de relatie met de reconstructie van het gladde deel van de mediumcoëfficiënten (snelheidsanalyse).

In hoofdstuk 4 beschouwen we akoestische media waarbij we er vanuit gaan dat niet is voldaan aan de aanname die in de literatuur bekend staat als reistijdinjectiviteit. We gaan er vanuit dat bronnen en ontvangers zich op een open deel van de rand van het medium bevinden. De data worden gemodelleerd met de Born-benadering. We laten zien dat in sommige gevallen inversie voor de singuliere mediumverstoring niet mogelijk is. Echter, generiek bestaat de inverse wel. We lichten dit toe met voorbeelden.

Dankwoord

De afgelopen vier jaar heb ik met veel plezier aan dit proefschrift gewerkt. Nu het af is wil ik graag iedereen bedanken die aan de totstandkoming ervan heeft bijgedragen.

Allereerst zijn dat natuurlijk mijn begeleiders, Hans Duistermaat van het Mathematisch Instituut en Fons ten Kroode van Shell Research. Hun vragen en ideeën zijn voortdurend een belangrijke inspiratiebron voor dit onderzoek geweest. Ook heb ik veel opgestoken van de privécolleges die ik regelmatig kreeg van Hans Duistermaat. Ik bedank ook Martijn de Hoop van de Colorado School of Mines, met wie ik hoofdstuk 3 heb geschreven. Het was een leuke en leerzame ervaring om een tijd in Colorado te werken.

I would like to thank the reading committee, consisting of Odo Diekmann, Martijn de Hoop, Dirk Smit and Bill Symes for their useful comments on the manuscript.

Uiteraard wil ik Shell Research bedanken voor het financieren van mijn aanstelling. Ik bedank het Center for Wave Phenomena van de Colorado School of Mines, dat mij uitgenodigd heeft om enige tijd in Colorado te werken. NWO wil ik bedanken voor het verstrekken van twee SIR reisbeurzen om een zomerschool respectievelijk een conferentie te bezoeken, en de Europese Unie voor het ondersteunen van deelname aan een zomerschool.

Ik wil mijn collega's van het Mathematisch Instituut bedanken voor de prettige werksfeer, in het bijzonder mijn kamergenoten Alexa, Martijn, Giovanna en Michael. Veel mensen hebben eraan bijgedragen dat ik ook naast mijn AIO-schap een leuke tijd heb beleefd. Graag wil ik daarvoor mijn familie en vrienden bedanken, met name Emiel, Geertje, Lonneke en Therese en de mensen van tafeltennisvereniging Otio. Tot slot wil ik Ingrid bedanken voor alles wat we de laatste jaren samen gedeeld hebben en voor haar steun, ook als ik vanwege mijn werk een tijd in het buitenland was.

Curriculum Vitae

Chris Stolk is op 24 december 1973 in Bunnik geboren. Van 1986 tot 1992 bezocht hij het Christelijk Gymnasium te Utrecht. In de laatste twee jaren hiervan deed hij mee aan de Nationale Wis- en Natuurkunde Olympiades, waar hij onder andere een eerste en twee tweede plaatsen behaalde. Deelname aan de Internationale Wis- en Natuurkunde Olympiades leidde tot drie zilveren medailles en een bronzen medaille.

Van september 1992 tot december 1996 volgde hij aan de Universiteit Utrecht de opleidingen Wiskunde en Theoretische Natuurkunde. In het studiejaar 1994/1995 verbleef hij een jaar als uitwisselingsstudent in Berkeley, waar hij colleges over quantumveldentheorie en algemene relativiteitsleer volgde. Daarna schreef hij een afstudeerscriptie op het gebied van de quantumveldentheorie bij prof. B.Q.P.J. de Wit. In januari 1997 resp. december 1996 studeerde hij cum laude af in de Theoretische Natuurkunde en in de Wiskunde.

Vanaf januari 1997 werkte hij aan dit proefschrift als AIO bij het Mathematisch Instituut van de Universiteit Utrecht. In die tijd bracht hij twee langere werkbezoeken aan de Colorado School of Mines. Hij nam deel aan de Studiegroep Wiskunde met de Industrie, bezocht zomerscholen over mathematische geofysica en inverse problemen en presenteerde zijn werk op verscheidene conferenties.

Vanaf 1 januari 2001 zal hij als postdoc werkzaam zijn aan de Rice University in Houston.