Born Inversion in Realistic Backgrounds by Means of Recursive Green's Functions

T.J. Moser [1,2], M. Biryulina[1], and G. Ryzhikov[1]

[1] Institute of Solid Earth Physics, University of Bergen, Norway
[2] now at: Geophysikalisches Institut, Karlsruhe, Germany

Abstract. The commonly applied methods for seismic inversion are based on some drastic assumptions regarding the known background or macro-velocity model and the data acquisition, that limit their applicability in geologies of realistic complexity or to realistic, noisy and incomplete, data sets. The background is usually assumed smooth, often to such a degree that the wave field can be described by simple ray theory without caustics or multipathing; the data are assumed to be complete and noise-free. Correspondingly, the algorithms for ray-based Green's functions are, until now, mostly developed for smooth media. To be able to image in realistic backgrounds and with realistic data sets, the assumptions have to be weakened. This must be done on two fronts: the imaging formula and the Green's functions. A new, generalized, imaging formula has been developed that takes into account that real data are incomplete, noisy and have a limited frequency band. A new approach for Green's functions allows the backgrounds in the inversion to be non-smooth, and accounts for reflected and transmitted ray fields by organizing the ray tracing recursively. Combined, the two approaches allow a systematical target-oriented inversion, in which upper parts of the Earth model are assumed known and fixed, and the attention is concentrated on important details below. The new imaging formula, together with the realistic Green's functions, has been successfully applied on the imaging of a complicated horst structure from the North sea.

Keywords. Born inversion, raytracing, Green's functions

1 Introduction

A seismic inversion in a realistic background remains a challenge even for 2D problems, because of the absence of any a priori symmetry. An advanced inversion formula, reducing inversion to a sort of migration, has been suggested by Beylkin ([2], [3]), and applied with relative success ([7], and others). Nevertheless, this formula has a limited applicability in areas with a realistic complexity, or for realistic seismic data acquisitions. The inversion is based on the generalized Radon transform, which supposes that for a given background and a given source/receiver acquisition, there exists a one-to-one correspondence between a record sample and the parameters of an isochron crossing a background point. Even in a layered background this supposition fails due to existence of up- and downgoing rays, crossing the background point. Besides, the inversion formula is not valid for incomplete data (for instance, reflection data from a narrow aperture acquisition), also it is hard to adapt the formula for data from a few source/receiver pairs, noisy data, narrow frequency band signal, etc. Also, its validity depends strictly on the validity of ray theory, which, despite the computational speed of ray tracing, is rather limited. That is why it is preferable to deal with an inversion based on a general wave field evaluation, rather than explicitly on ray theory. The inversion formula suggested by Ryzhikov and Troyan ([12], [13], [14]) allows to perform "wave field + Born" -, or "wave field + linearized response"- inversion. A detailed derivation is included in the appendix.

Based on this more generalized inversion scheme, this paper presents an attempt to image a complicated horst structure in the North Sea. The background, or the macro-velocity model, is allowed to be non-smooth in this numerical experiment. The Green's functions are provided by the recursive seismic ray modeling scheme ([9], [10], [11]), which is summarized below. To show the viability of the proposed imaging scheme, the inversion of a synthetic data set is presented in this paper; real data are inverted by the same scheme in [10].

2 Inversion

Let us try to summarize why so elegant a concept on the interpretation of dynamical seismic data as the generalized Radon transform fails, when it is applied to raw real data. Mathematically, this is caused first of all by the background's Green's function, appearing in the Lippmann-Schwinger equation. To deal with data in terms of the first Born approximation ('linear'/'single-scattering'/'single-diffracting'), we have to start from the background's Green's function being far from the ray-theoretical one, otherwise we have to include many, or even an infinite number, of terms in the Born series (e.g. a perturbation like a full layer can be represented only by an infinite number of terms). If we want to maintain the concept of an inversion of the generalized Radon transform, then the $2{\times}2$ Jacobian of the mapping from local parameters describing an isochron to a generic source/receiver location should exist. Apart from that, there should be a one-to-one correspondence between the pair 'sample of a source-receiver pair record' - 'location + orientation of isochron' (for a 3D problem these are 3 parameters of a plane, tangent to an isochron). These requirements have quite drastic consequences for the inversion. Namely, they imply that:

a) the reference medium should be fairly simple: a signal propagation just in terms of rays (every sample has its own 'travel time'; every unit volume element of the medium provides a piece of information that, for a given source-receiver pair record, is associated just with the pair of rays 'source - medium point - receiver'), no caustics, no shadow zones, locally plane isochrons, etc. (which is not valid even for a stratified reference medium or a medium with a constant velocity gradient). Note that the representation of a (generally curved) isochron by a plane is only locally valid, in a vicinity of the touching point of the tangent plane to the isochron (e.g. for a small object in a far field zone of a near to homogeneous reference medium);

b) the effect of the reference medium perturbation should be linear, which

corresponds to a weak and single scattering/diffraction;
c) for a given source the spatial parameters of the receiver should run over the full sphere surrounding the target of the imaging. This may be realizable in technical devices (e.g. tomograph), but it is unfeasible for lots of remote sensing, geophysical, and technical problems, that naturally suffer from a finite aperture. Any attempt to complete the data by an interpolation/extrapolation in data space indirectly induces an a priori representation of reconstructed medium parameters (e.g. a smoothing of data is associated with a corresponding smoothing of unknown media parameters - hard to analyze).

Mathematical problems arise already when the ray-theoretical representation is fairly good, but apart from that, we have to:
1. introduce a non-smooth reference model (e.g. layered medium);
2. use a sounding signal of a finite frequency band, i.e. a signal of finite time duration, not a delta impulse;
3. deal with data from acquisition involving a few sources: it is hard to avoid overlapping of data;
4. interprete data from a realistic experiment design.

Besides, as a rule real data are incomplete: they are not given in continuous-, but in digital form, and the set of receiver positions is finite, sparse or even random. Also the data are noisy, due to variety of known and unknown sources; this makes regularization difficult. But, of course, the main problems are generated by wave phenomena, which depend upon medium supposed to be known a priori and upon the ratios 'dominant wave length/perturbation size/source-object-receiver distances'. Very often such phenomena can not be described properly in terms of the ray theory; the latter is the base to treat dynamical data in terms of integrals over isochrons.

Therefore we follow another way, which is not so strongly associated with the kinematics of first arrivals. We start from realistic assumptions about the data, taking into account their strong incompleteness, i.e., their inadequacy to determine the unknown medium parameters. For example, for seismic exploration the data are represented by unstacked digitized finite-band noisy seismograms, recorded by a sparse and probably random net of sources and receivers. The strategy of our approach is tied to dealing with such segments of data, which can be decoded properly. It means that these segments should be described by a linearized forward model with an accuracy induced by the value of data noise, - the latter is caused by the physical registration channel and unidentified sources. The proper input for the inversion are data residuals: observed data reduced with the wave field evaluated for unperturbed background. Note here, that such a reduced data set can be easily obtained directly in a problem of reconstruction of inclusions in a laterally uniform background or in a problem of seismic monitoring: these are just differential data. The inverse problem is posed as an optimization one: to find the perturbation of a background, that provides the best fitting of corresponding synthetic (linear) responses with input data for all source-receiver pairs. The linear response, caused by a volume unit perturbation consists of: evaluation of the wave field in a background generated by a source (incoming wave field); application of an operator of interaction - the result can be interpreted as a secondary source; evaluation of the wave field (in the background), generated by the secondary source (outgoing wave field). The operator of interaction consists of a Fréchet-derivative of a wave operator with respect to medium parameters; e.g. for the acoustic wave equation this operator is the well-known second time derivative, other examples of Fréchet derivatives/operators of interaction for an arbitrary background are given in [12]. The actual inversion formula is the result of a simplified solution of the optimization problem, and provides the local inversion, just as Beylkin's formula does: to get the result of inversion in a given volume, we need to collect just linear responses of data, induced by unit perturbation in this volume only. It allows to parallelize the inverse problem very efficiently. The formula can be interpreted as follows: to reconstruct the value of perturbation in a given volume, it is necessary to evaluate a correlation coefficient between data residuals for all (relevant) source-receiver pairs and all linear responses induced by a unit perturbation in the given volume element. This correlation coefficient contains a factor, being a ratio of an average amplitude of data residuals and average amplitude of linear responses.

A short summary of our approach is following (a detailed derivation is given in the appendix):
1. No need to deal with ray/isochron representation: the inversion is expressed in terms of in- and outgoing wave fields in a reference medium, no matter how they are evaluated (e.g. by finite differences). This allows an essential reduction of errors, that is to be accounted for by a proper forward model, being a base for inversion. When the approximation ray + Born is adequate to real phenomena, our forward model is the same as the conventional one, except that we take into account directly the effect of registration channel (to be realized a channel should have a finite frequency band).
2. The algebraization/discretization of the problem is straightforward - it corresponds to finite number of records with digital data.
3. The problem is then reduced to an optimization one, that allows to treat ill-posed problems properly.

It is evident that strong perturbations generally can not be expressed in terms of the Born/generalized Born finite series. For example, if the perturbation of a stratified reference medium is represented by a thick layer, it leads to a phase shift with respect to the propagation in the reference medium. This shift can not be described properly by the first term of the Born series, nor by any finite number of terms. However, there are data segments that can be represented well as linear response. For example, when an overburden is known and our problem is to reconstruct an unknown interface of a bedrock. Although the linear response is not valid for the transmitted data, it is then still a good approximation for the short-offset reflected data. In our numerical experiments we supposed that these data can be simulated with the Born-diffractors located in a vicinity of the interface [4]. The solution of a more general 3D nonlinear inverse problem was shown in [14].

3 Recursive construction of Green's functions

Like the one presented in this paper, most inversion schemes heavily rely on the possibility of a computationally fast forward modeling of the wave propagation, or the availability of Green's functions for the wave equation, in a presumed known Earth model (i.e. reference- or background model, see appendix equations [A4] and [A5]). Although the proposed inversion procedure is valid for any type of forward modeling, ray theory still provides by far the fastest algorithms. The ray-theoretical representation of the Green's function is $G(\boldsymbol{x}, \boldsymbol{y}, \omega) = \sum_k A_k(\boldsymbol{x}, \boldsymbol{y}) \exp i\omega T_k(\boldsymbol{x}, \boldsymbol{y})$, where $T_k(\boldsymbol{x}, \boldsymbol{y})$ is the traveltime between two locations $\boldsymbol{x}$ and $\boldsymbol{y}$, $A_k(\boldsymbol{x}, \boldsymbol{y})$ the corresponding (possibly complex-valued) amplitude and the summation over k accounts for multipathing. In recent years, schemes have been developed for ray-based isochron- or wavefront tracing in smooth media without reflecting interfaces ([5], [6], [15], [16], [17]). They construct the Green's function for one fixed point $\boldsymbol{y}$ (for instance a source or receiver point) and for $\boldsymbol{x}$ running over the model, by propagating a set of rays from $\boldsymbol{y}$ and resample the rays at each computed new isochron, in order to achieve a uniform coverage of the model by rays. In such a way receiver captures are easily detected and each time the wavefront passes a receiver point the ray quantities necessary for a local evaluation of the wave field (arrival time, amplitude, phase shift) can be easily assessed by a simple interpolation from neighboring rays. By allowing the wave front to fold over itself at caustics, the algorithms properly take multivaluedness into account. The two-point raytracing problem is thereby eliminated, and the whole model is filled with (possibly multivalued) ray arrivals.

The Born approximation for scattering by an obstacle is valid under the condition that the magnitude of the scattered wave, measured in some norm, is much smaller than that of the incoming- or reference wave. This implies that it is a weak-scattering- and low-frequency approximation (refer to appendix equation [A10]), contrary to the ray-based Green's functions, which are high-frequent approximate. A successful application of Green's functions in Born inversion therefore depends on how small the scatterer is, in other words, the allowed degree of complexity of the background; the more realistic the background, the smaller is the unknown scatterer and the better is the image. Ray-based Green's functions in smooth media impose severe restrictions in this respect, for a variety of reasons. When a smooth background is to be improved in order to approach a non-smooth reality (according to criteria not discussed in this paper), the ray fields become more and more multivalued. As a result, this implies that the computation time, for the Green's functions as well as for the imaging, increases accordingly. Also, along the caustics the ray-amplitude is infinite, so that the image quality is distorted (often only thanks to numerical inaccuracies a complete break-down is prevented). Next, for a sufficient accuracy the integration step-size along the rays has to decrease, increasing again the CPU. Finally, ray theory is not valid any more in smooth, but rapidly varying media, and does not represent any more the actual (finite-frequency) wave field. The effects of such fields on the image quality is shown in the section 4.

All these considerations motivate the introduction of discontinuities in the background and modify the smooth medium's Green's functions algorithms to include reflections and transmissions. One disadvantage of non-smoothness is that, unless diffractions are taken into account (which is expensive), it can introduce shadow zones in the Green's functions. However, we think that, as long as the Green's functions from all data points together provide a sufficient coverage of the target region, imaging with shadow zones, i.e. zero amplitudes, is to be preferred to imaging with caustics, i.e. infinite amplitudes. Apart from diffractions, the ray- and wave field is already considerably more complicated in presence of discontinuities than in smooth models, due to multiple reflections, which themselves may be multivalued again, not only because of smooth model variations, but also because of possible curvature of the discontinuities. The organization and storage of all these different phases can cause formidable difficulties.

Table 1. Synthetic velocity model

Layer	Velocity (km/s)
1	1.5
2	2.2
3	2.9
4	2.4
5	2.9
6	3.4
7	4.2

Table 2. Brage velocity model

Layer	Name	Velocity (km/s)
1	water	1.478
2	upper sediments	1.900
3	Tertiary I	2.020
4	Tertiary II	1.815
5	Cretacious	2.680
6	Balder	2.350
7	Shetland	2.950
8	Draupne	2.900
9	Fensfjord	3.464
10	Brent	3.100
11	fault zone	2.800
12	Shetland	3.100
13	Statfjord	3.500

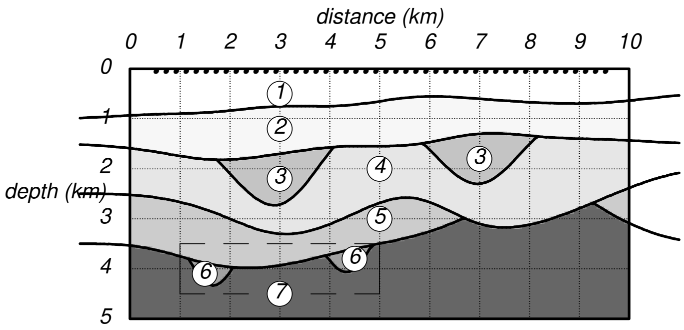

Fig. 1. Synthetic velocity model. The labels correspond to layers tabulated in Table 1. The gray scale represents velocities. The box bounds the target region. The dots along the surface denote the shot/receiver locations.

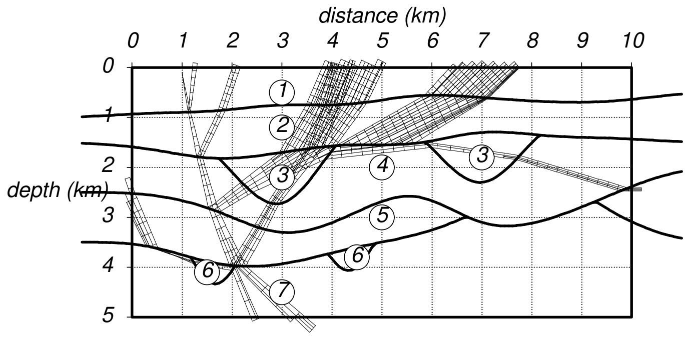

Fig. 2. Recursive raytracing of one ray cell in the synthetic model. Transmissions and first order reflections are shown.

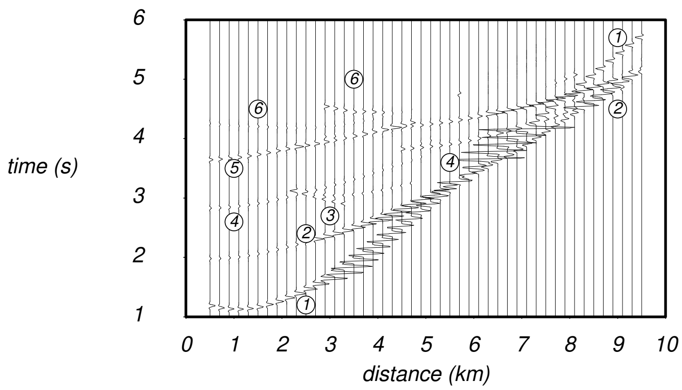

Fig. 3. Synthtic section for a shot at 1.0km in the synthetic model, computed by recursive raytracing. Only first order reflections are shown. The labels refer to the layer in which the reflection occurred. The arrivals labeled 5 and 6 are used for the images in Figures 4 to 7. Wavelet: 10 Hz Gabor.

A recursive treatment ([9], [10], [11]) solves these problems by reversing the order of the computations; instead of operating with wave fronts and isochrons, it works with only one ray cell at a time, defined by two (in 2D) and three (in 3D) neighboring rays and two successive isochrons. This ray cell is originated at a source point or at an initial surface, and propagated all the way through the medium until it meets some termination criterion. At interfaces, the ray cell splits into a reflected and a transmitted cell, that both continue their own way, independently from each other. Both cells possibly hit new interfaces, each time generating new offsprings, that behave similarly to the original cell. The tree structure of ray segments, that thus appears, is conveniently handled by recursion. The storage requirement for the tree is negligible, because it depends only logarithmically on the number of subdivisions in the family of one initial cell. One particular advantage of recursion is that all relevant phases on a seismic section are generated automatically, without human interaction, together with all information necessary to analyze or select them; this advantage has been exploited in phase identification in vertical seismic profiling [10]. A detailed description and evaluation of the recursion algorithm is given in [10], both in 2D and 3D. Also in [10], an approach to include edge diffractions in the recursion is presented; diffractions fill in shadow zones, but may be expensive.

4 Numerical tests

The numerical tests have been performed on two models, a synthetic model, to illustrate the statements in the previous sections, and a real model, the Brage oil field from the Northern part of the North Sea, which is in production by Norsk Hydro. In both cases synthetic data are used, in the synthetic model provided by raytracing, in order to isolate possible imaging artefacts, in the Brage model provided by a finite-difference solution of the wave equation, in order to test the viability of the imaging formula, without interference of typical problems of real data (like noise, multiples, 3D effects, elastic effects, or inaccurate information on the overburden). Real Brage data are inverted with the same imaging technique and with non-smooth medium's Green's functions algorithms in [10].

The synthetic model is shown in Figure 1 and its velocities (constant per layer) are tabulated in Table 1. Special attention is given to the 'channel' structures, labeled '6' is Figure 1. Figure 2 illustrates the recursive raytracing of one cell in the

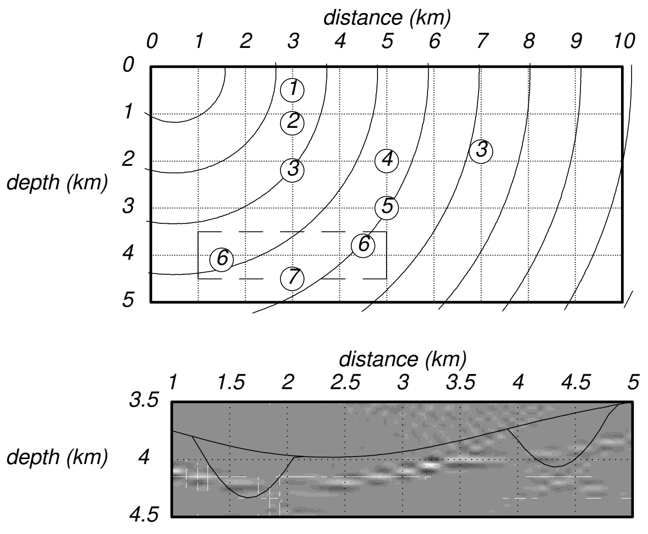

Fig. 4. Isochrons of the Green's function in a constant background (above). Rays are not shown. The reconstructed image in the target region is shown below (gray scale: perturbation of squared slowness on original model, thin lines: original interfaces).

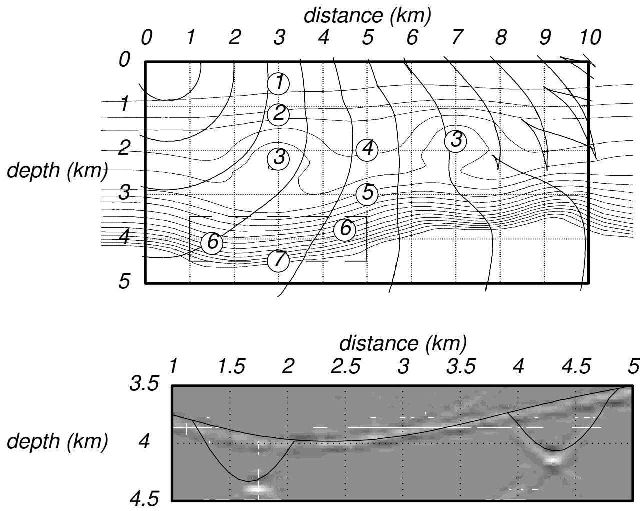

Fig. 5. Isochrons of the Green's function in a slowly varying background (above) and image belonging to it (below).

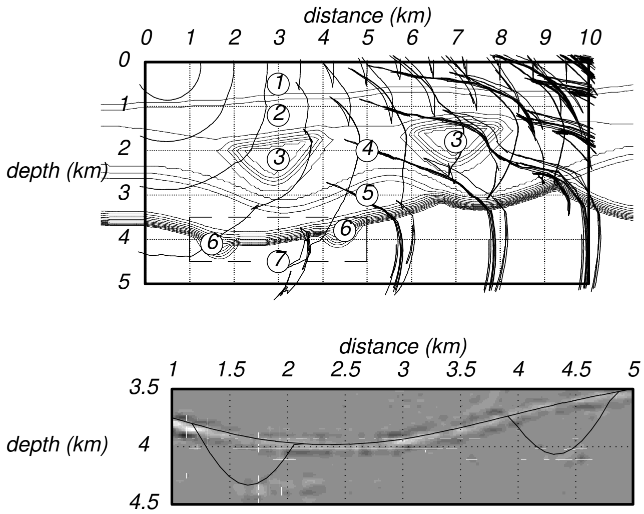

Fig. 6. Isochrons of the Green's function in a rapidly varying background (above) and image belonging to it (below).

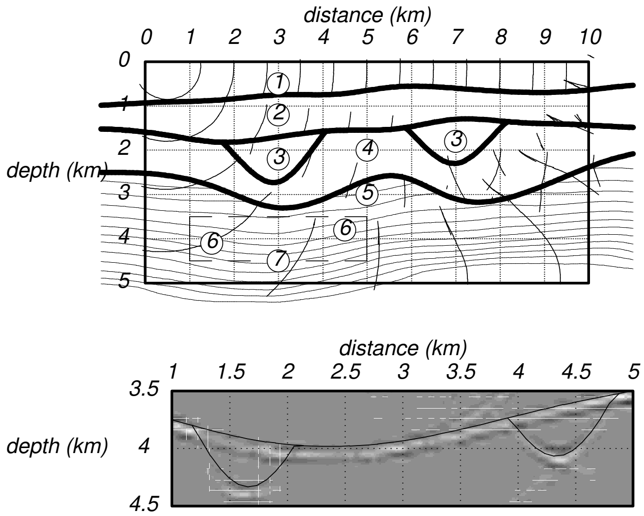

Fig. 7. Isochrons of the Green's function in a non-smooth background (above) and image belonging to it (below).

synthetic model. For a complete coverage of the model, and to obtain the shot record in Figure 3, ray cells are initiated in al directions from the source point. In Figure 2, one sees how, already from one initial cell, a quite complicated tree of ray segments, reflected and transmitted, emanates. Thanks to the recursive organization, it is possible to accumulate the ray history along the ray and to distinguish the phases in Figure 3 by the layer in which they were reflected. 46 locations act as sources and receivers in the synthetic model, from where Green's functions into the model are calculated. 32 receiver locations symmetrically around each source point are involved in the imaging. Figures 4 to 7 show the imaging with the formula [A19] (appendix), for a sequence of reference models that are increasingly complex and close to the actual model. In Figures 4 to 6, the reference models are smooth. It is clearly visible how the Green's functions become more and more multivalued, developing caustics and finally distorting the image quality (Figure 6). Figure 7 shows the imaging with a non-smooth reference field. Shadow zones appear, but the Green's functions are much less multivalued and therefore realistic, and the image is correspondingly better (and cheaper to compute).

An interpreted depth section of the Brage model is shown in Figure 8 and its velocities are tabulated in Table 2. It contains lots of diffracting edges, so that the ray field is disconnected and exhibits shadow zones (Figure 9, for one initial cell, and Figure 10 a shot record). Yet, a comparison with finite-differences, which is supposed to provide the complete wave field (at the expense of a much higher CPU) shows agreement to a high degree in arrival times, amplitudes and phase shifts (Figure 10 against 11). The data acquisition used for the imaging is tabulated in Table 3. The non-smooth reference model has been chosen equal to the velocity

Table 3. Data acquisition in Brage

Number of shots	400
Number of receivers/shot	120
First shot location (km)	-10.0
Shot location increment (km)	+0.0375
First receiver offset (km)	+0.1375
Receiver offset increment (km)	+0.025
Shot/receiver depth (km)	+0.008
Sampling rate (s)	+0.002

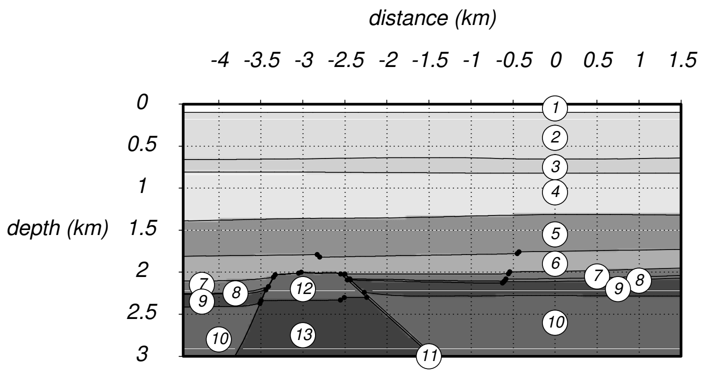

Fig. 8. The Brage velocity model. The labels correspond to the stratigraphic units tabulated in Table 2. The gray scale represents velocities. The black dots denote diffracting edges.

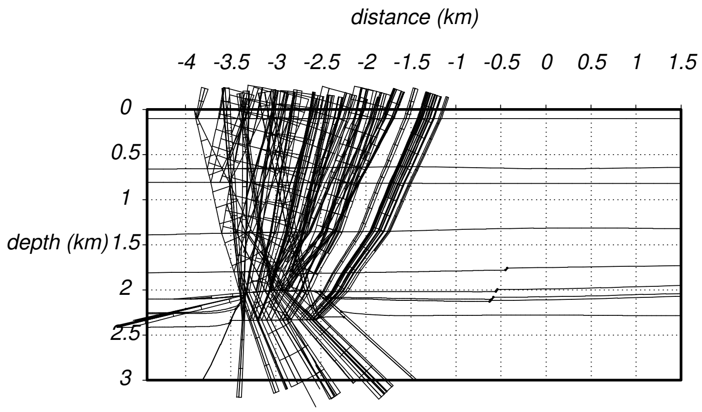

Fig. 9. Recursive raytracing of one ray cell in the Brage model. Transmissions and first order reflections.

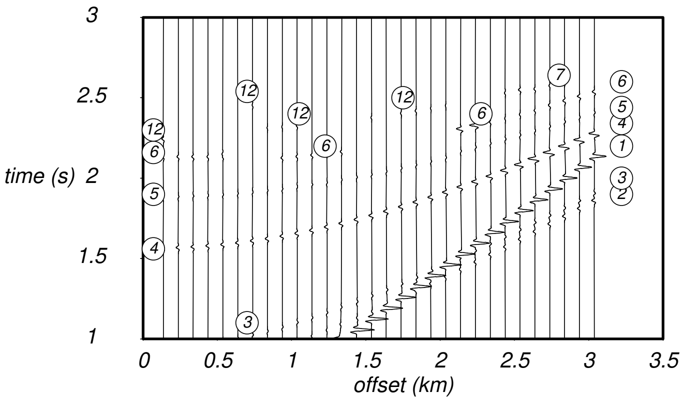

Fig. 10. Synthetic section for a shot in the Brage model at -3.9km, computed by recursive raytracing. The labels refer to reflections on the bottom of layers in Figure 8. Wavelet: 30 Hz Gabor.

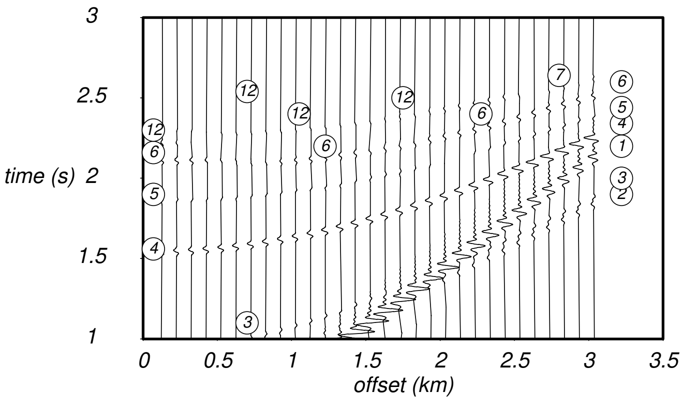

Fig. 11. Finite-difference section for a shot in the Brage model at -3.9km. The labels are taken from Figure 10. The arrivals labeled 12 are used in the inversion to obtain the image of Figure 13. Wavelet: 30 Hz zero-phase Ricker.

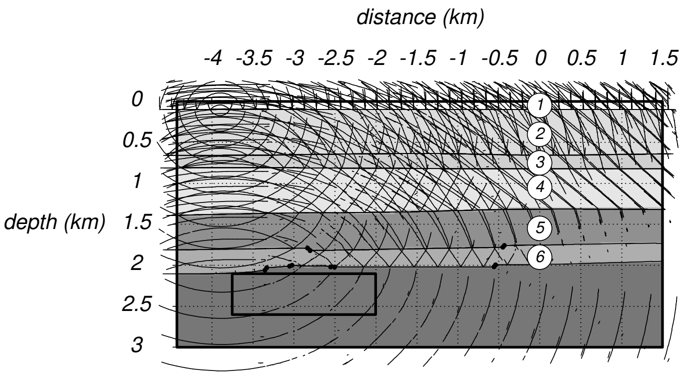

Fig. 12. Non-smooth reference model for inversion. The target is bounded by the rectangular box. The gray scale represents velocities. The thin lines denote the isochrons of the Green's function for a source at -3.9 km.

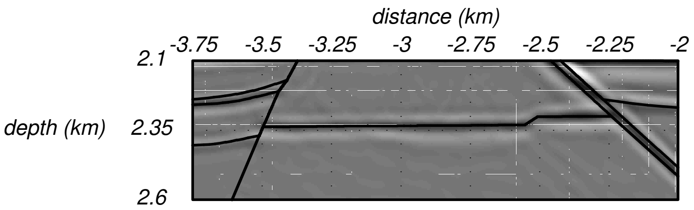

Fig. 13. Image of the horst interior. The grey scale denotes the reconstructed perturbation of the squared slowness on the reference model given in Figure 12. The thin lines represent the geometry of the Brage-model, in which the (finite-difference) synthetic data have been generated (see Figure 11).

model from the interpreted depth section (Figure 8), as far as the overburden is concerned (indicated by indices 1 to 6), and chosen equal to a constant representative velocity in the lower part of the model. The target zone is bounded by the rectangular box in Figure 12 and contains the most important, interior, part of the horst. Also shown are the isochrons of the Green's functions for an exemplary source point at -3.9km. The image derived from the finite-difference data is displayed in Figure 13. The geometry of the interfaces is very accurately reconstructed. Dynamical information is available as well, in the form of a reflectivity measure; the right fault has a sharp contrast, the fault on the left almost no contrast.

Acknowledgements

The authors acknowledge financial support from the Norwegian Scientific Council for the EU-project 3D Asymptotic Seismic Imaging, NRF No 30995. We are grateful to Dr J. Pajchel from Norsk Hydro for his invaluable contributions to the work and for running many numerical tests. The first author is grateful to the Alexander-von-Humboldt Foundation for financial support.

References

[1] Beydoun, W.B., and Mendes, M., 1989. Elastic ray-Born l_2-migration/inversion, *Geophysical Journal*, **97**, 151-160.

[2] Beylkin, G., 1985. Imaging of discontinuities in the inverse scattering problem by inversion of a causal generalized Radon transform, *J. Math. Phys.*, **26**, 99-108.

[3] Beylkin, G., and Burridge, R., 1990. Linearized inverse scattering problem in acoustics and elasticity, Wave Motion, **12**, 15-52.

[4] Biryulina, M., and Ryzhikov, G. 1995. Rytov-Born decomposition in 3D reflection seismics, 57th Annual EAGE Meeting Glasgow, Expanded Abstracts.

[5] Lambaré, G., Lucio, P.S., and Hanyga, A., 1994. 2D fast asymptotic Green's function, 56th Annual EAEG Meeting Vienna, Expanded Abstracts.

[6] Lucio, P.S., Lambaré, G., and Hanyga, A., 1995. 3D multivalued travel time and amplitude maps, 57th Annual EAEG Meeting Glasgow, Expanded Abstracts.

[7] Miller, D., Oristaglio, M., and Beylkin, G., 1987. A new slant on seismic imaging: Migration and integral geometry, *Geophysics*, **52**, 943-964.

[8] Moser, T.J., 1995. Inversion in a non-smooth background, 57th Annual EAEG Meeting Glasgow, Expanded Abstracts.

[9] Moser, T.J., and Pajchel, J., 1996. 3D seismic profiling by means of recursion, 58th Annual EAGE meeting Amsterdam, Expanded Abstracts.

[10] Moser, T.J., and Pajchel, J., 1996. Recursive seismic ray modeling: applications in inversion and VSP, submitted to *Geophysical Prospecting*.

[11] Pajchel, J., and Moser, T.J., 1995. Recursive cell raytracing, 51th Annual EAEG Meeting Glasgow, Expanded Abstracts.

[12] Ryzhikov, G. and Troyan, V., 1992. 3D diffraction tomography. Part 1: Construction and interpretation of tomography functionals, in *Expanded Abstracts, Russian - Norwegian Oil Exploration Workshop II*, Voss.

[13] Ryzhikov, G. and Troyan, V., 1992. 3D diffraction tomography. Part 2: Reconstruction algorithm with statistical regularization, in *Expanded Abstracts, Russian - Norwegian Oil Exploration Workshop II*, Voss.

[14] Ryzhikov, G., Biryulina, M., and Hanyga, A., 1995. 3D nonlinear inversion by entropy of image contrast optimization, *Nonlinear Processes in Geophysics*, **2**, 228-240.

[15] Sun, Y., 1993. Computation of 2-D multivalued arrival travel time fields by an interpolative shooting method, 63th Annual SEG Meeting, Expanded Abstracts.

[16] Vinje, V., Iversen, E., Gjøystdal, H., and Åstebøl, K., 1994. Estimation of multivalued arrivals in 3D models using wavefront construction, 56th Annual EAEG Meeting Vienna, Expanded Abstracts.

[17] Vinje, V., Iversen, E., and Gjøystdal, H., 1993. Traveltime and amplitude estimation using wavefront construction, *Geophysics*, **58**, 1157-1166.

Appendix - Imaging formula

Scattering phenomena for the scalar wave equation

$$\Delta\phi(\boldsymbol{x},\ t) - u^2(\boldsymbol{x})\,\frac{\partial^2\phi(\boldsymbol{x},\ t)}{\partial t^2} = S(\boldsymbol{x},\ t) \qquad \text{[A1]}$$

can be analyzed by separating the squared slowness $u^2(\boldsymbol{x})$ in a reference field and a perturbation on it:

$$u^2(\boldsymbol{x}) = u_0^2(\boldsymbol{x}) + \varepsilon u_1^2(\boldsymbol{x})\ , \qquad \text{[A2]}$$

and expanding the wave field $\phi(\boldsymbol{x},\ t)$ in a Born series

$$\phi(\boldsymbol{x},\ t) = \sum_{n=0}^{\infty} \varepsilon^n \phi_n(\boldsymbol{x},\ t)\ . \qquad \text{[A3]}$$

Inserting [A2] and [A3] in [A1] and collecting the ε^0 terms gives the wave propagation in the reference model:

$$\Delta\phi_0(\boldsymbol{x},\ t) - u_0^2(\boldsymbol{x})\,\frac{\partial^2\phi_0(\boldsymbol{x},\ t)}{\partial t^2} = S(\boldsymbol{x},\ t)\ , \qquad \text{[A4]}$$

which is formally solved by the Green's function $G(\boldsymbol{x},\ \boldsymbol{y},\ t,\ \tau)$:

$$\phi_0(\boldsymbol{x},\ t) = \int_M \int_{-\infty}^{+\infty} G(\boldsymbol{x},\ \boldsymbol{y},\ t,\ \tau) S(\boldsymbol{y},\ \tau)\ \mathrm{d}\boldsymbol{y}\ \mathrm{d}\tau\ . \qquad \text{[A5]}$$

In [A5], M denotes the spatial support of the integrand; the time integration extends from $-\infty$ to $+\infty$, but in practice over the length of a seismic trace. In [A1], [A4], and [A5], $S(\boldsymbol{x},\ t)$ denotes the source density. For a point source, located at $\boldsymbol{x}_s$,

$$S(\boldsymbol{x},\ t) = \delta(\boldsymbol{x} - \boldsymbol{x}_s) s(t) \qquad \text{[A6]}$$

and

$$\phi_0(\boldsymbol{x},\ \boldsymbol{x}_s,\ t) = \int_{-\infty}^{+\infty} G(\boldsymbol{x},\ \boldsymbol{x}_s,\ t,\ \tau) s(\tau)\ \mathrm{d}\tau\ . \qquad \text{[A7]}$$

Collecting the ε^1 terms results in the equation for the first-order scattering:

$$\Delta\phi_1(\boldsymbol{x},\ t) - u_0^2(\boldsymbol{x})\,\frac{\partial^2\phi_1(\boldsymbol{x},\ t)}{\partial t^2} = u_1^2(\boldsymbol{x})\,\frac{\partial^2\phi_0(\boldsymbol{x},\ t)}{\partial t^2}\ , \qquad \text{[A8]}$$

which is again formally solved by the Green's function

$$\phi_1(\boldsymbol{x},\ t) = \int_M \int_{-\infty}^{+\infty} G(\boldsymbol{x},\ \boldsymbol{y},\ t,\ \tau) u_1^2(\boldsymbol{y})\,\frac{\partial^2\phi_0(\boldsymbol{y},\ \tau)}{\partial \tau^2}\ \mathrm{d}\boldsymbol{y}\ \mathrm{d}\tau\ . \qquad \text{[A9]}$$

The frequency domain expression for [A9] reads

$$\phi_1(\boldsymbol{x},\ \omega) = -\omega^2 \int_M G(\boldsymbol{x},\ \boldsymbol{y},\ \omega) u_1^2(\boldsymbol{y}) \phi_0(\boldsymbol{x},\ \omega)\ \mathrm{d}\boldsymbol{y}\ . \qquad \text{[A10]}$$

For a reference wave excited by a point source (see [A7])

$$\phi_1(\boldsymbol{x},\, \boldsymbol{x}_s,\, t) = \int_M \int_{-\infty}^{+\infty} G(\boldsymbol{x},\, \boldsymbol{y},\, t,\, \tau) u_1^2(\boldsymbol{y}) \frac{\partial^2 \phi_0(\boldsymbol{y},\, \boldsymbol{x}_s,\, \tau)}{\partial \tau^2}\, \mathrm{d}\boldsymbol{y}\, \mathrm{d}\tau\ . \qquad [A11]$$

The image is derived from the least-square misfit between the first-order scattered wave and the observed wave ϕ^{obs}:

$$F[u^2] = \frac{1}{2} \sum_{r,s} \int_{-\infty}^{+\infty} (\phi_0(\boldsymbol{x}_r,\, \boldsymbol{x}_s,\, t) + \phi_1(\boldsymbol{x}_r,\, \boldsymbol{x}_s,\, t)[u^2] - \phi^{obs}(\boldsymbol{x}_r,\, \boldsymbol{x}_s,\, t))^2\, \mathrm{d}t\ . [A12]$$

In [A12], the summation is over all sources $\boldsymbol{x}_s$ and receivers $\boldsymbol{x}_r$ available in the data acquisition. Here, and in what follows, we put $\varepsilon = 1$. $F[u^2]$ is a functional, that assigns a real number to each squared slowness distribution $u^2(\boldsymbol{x})$; the dependence on u^2 is denoted by square brackets. Its gradient with respect to u^2, $\nabla_{u^2} F[u^2]$, is a function of a spatial location $\boldsymbol{y}$. In the first Born approximation, the scattered wave ϕ_1 depends linearly on u^2, and $F[u^2]$ has one minimum, which follows from

$$\nabla_{u^2} F[u^2](\boldsymbol{y}) = 0\ . \qquad [A13]$$

An expression for $\nabla_{u^2} F[u^2]$ is obtained by differentiating [A12]:

$$\begin{aligned} \nabla_{u^2} F[u^2](\boldsymbol{y}) = \sum_{r,s} \int_{-\infty}^{+\infty} (\phi_0(\boldsymbol{x}_r,\, \boldsymbol{x}_s,\, t) + \phi_1(\boldsymbol{x}_r,\, \boldsymbol{x}_s,\, t)[u^2] \\ - \phi^{obs}(\boldsymbol{x}_r,\, \boldsymbol{x}_s,\, t)) \nabla_{u^2} \phi_1(\boldsymbol{x}_r,\, \boldsymbol{x}_s,\, t)[u^2](\boldsymbol{y})\, \mathrm{d}t\ . \end{aligned} \qquad [A14]$$

In [A14], $\nabla_{u^2}\phi_1(\boldsymbol{x}_r,\, \boldsymbol{x}_s,\, t)[u^2](\boldsymbol{y})$ is a function denoting the gradient of the scattered wave with respect to the squared slowness perturbation. It can be expressed with help of [A11]:

$$\begin{aligned} I(\boldsymbol{x}_r,\, \boldsymbol{x}_s,\, \boldsymbol{y},\, t) \equiv \nabla_{u^2}\phi_1(\boldsymbol{x}_r,\, \boldsymbol{x}_s,\, t)[u^2](\boldsymbol{y}) = \\ = \int_{-\infty}^{+\infty} G(\boldsymbol{x}_r,\, \boldsymbol{y},\, t,\, \tau) \frac{\partial^2}{\partial \tau^2} \phi_0(\boldsymbol{y},\, \boldsymbol{x}_s,\, \tau)\, \mathrm{d}\tau\ . \end{aligned} \qquad [A15]$$

The solution to [A13] can be found by expanding $\nabla_{u^2} F[u^2]$ around the reference field u_0^2:

$$\nabla_{u^2} F[u^2](\boldsymbol{y}) = \nabla_{u^2} F[u_0^2](\boldsymbol{y}) + \int_M \nabla_{u^2}\nabla_{u^2} F[u_0^2](\boldsymbol{y},\, \boldsymbol{z})(u^2(\boldsymbol{z}) - u_0^2(\boldsymbol{z}))\, \mathrm{d}\boldsymbol{z}\ . \quad [A16]$$

The Hessian operator of $F[u^2]$ is represented by the integral kernel

$$\nabla_{u^2}\nabla_{u^2} F[u_0^2](\boldsymbol{y},\, \boldsymbol{z}) = \sum_{r,s} \int_{-\infty}^{+\infty} I(\boldsymbol{x}_r,\, \boldsymbol{x}_s,\, \boldsymbol{y},\, t) I(\boldsymbol{x}_r,\, \boldsymbol{x}_s,\, \boldsymbol{z},\, t)\, \mathrm{d}t\ . \qquad [A17]$$

Following [1], we approximate the Hessian by replacing it by its diagonal:

$$\nabla_{u^2}\nabla_{u^2}F[u_0^2](\boldsymbol{y},\, z) \sim \sum_{r,s} \int\limits_{-\infty}^{+\infty} I(\boldsymbol{x_r},\, \boldsymbol{x_s},\, \boldsymbol{y},\, t) I(\boldsymbol{x_r},\, \boldsymbol{x_s},\, z,\, t)\delta(\boldsymbol{y} - z)\, \mathrm{d}t\ . \quad [A18]$$

The image follows then from [A13] and [A16]:

$$u_1^2(\boldsymbol{y}) = u^2(\boldsymbol{y}) - u_0^2(\boldsymbol{y}) =$$
$$- \frac{\sum\limits_{r,s} \int\limits_{-\infty}^{+\infty} (\phi_0(\boldsymbol{x_r},\, \boldsymbol{x_s},\, t) - \phi^{obs}(\boldsymbol{x_r},\, \boldsymbol{x_s},\, t)) I(\boldsymbol{x_r},\, \boldsymbol{x_s},\, \boldsymbol{y},\, t)\, \mathrm{d}t}{\sum\limits_{r,s} \int\limits_{-\infty}^{+\infty} I(\boldsymbol{x_r},\, \boldsymbol{x_s},\, \boldsymbol{y},\, t)^2\, \mathrm{d}t}\ . \quad [A19]$$

The enumerator of [A19] is proportional to the data misfit and vanishes for a background equal to the real model. The denominator is referred to as information sensitivity and quantifies the possibility to reconstruct the image, given a certain data acquisition. The implementation of [A19] requires Green's functions to be calculated for all source and receiver locations towards the target region.