

**GEOLOGICA ULTRAIECTINA**

Mededelingen van de  
Faculteit Aardwetenschappen der  
Rijksuniversiteit te Utrecht

No. 97

**UPPER MANTLE STRUCTURE FROM  
BODY-WAVE CODA AND AMPLITUDES**

**FILIP NEELE**

GEOLOGICA ULTRAIECTINA

Mededelingen van de  
Faculteit Aardwetenschappen der  
Rijksuniversiteit te Utrecht

No. 97

**UPPER MANTLE STRUCTURE FROM  
BODY-WAVE CODA AND AMPLITUDES**

**FILIP NEELE**

CIP-DATA, KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Neele, Filippus Paulus.

Upper mantle structure from body-wave coda and amplitudes

/ Filippus Paulus Neele - Utrecht : Faculteit

Aardwetenschappen der Rijksuniversiteit Utrecht. -

(Geologica Ultraiectina, ISSN 0072-1026 ; no. 97)

Proefschrift Rijksuniversiteit Utrecht. - Met lit. opg. -

Met samenvatting in het Nederlands.

ISBN: 90-71577-51-1 geb.

Trefw.: P-golven / seismologie.

UPPER MANTLE STRUCTURE FROM  
BODY-WAVE CODA AND AMPLITUDES

BOVENMANTELSTRUCTUUR UIT CODA EN  
AMPLITUDES VAN RUIMTEGOLVEN

(met een samenvatting in het nederlands)

PROEFSCHRIFT

TER VERKRIJGING VAN DE GRAAD VAN DOCTOR AAN DE  
RIJKSUNIVERSITEIT TE UTRECHT OP GEZAG VAN DE  
RECTOR MAGNIFICUS PROF. DR. J.A. VAN GINKEL  
INGEVOLGE HET BESLUIT VAN HET COLLEGE VAN DECANEN  
IN HET OPENBAAR TE VERDEDIGEN OP DONDERDAG  
25 MAART 1993 DES NAMIDDAGS TE 12.45 UUR

DOOR

FILIPPUS PAULUS NEELE

GEBOREN OP 24 FEBRUARI 1965, TE OLDENZAAL

**PROMOTOR: PROF. DR N.J. VLAAR  
COPROMOTOR: DR R.K. SNIEDER**

**The research described in this thesis was carried out at the:**

**Department of Theoretical Geophysics  
Institute of Earth Sciences  
University of Utrecht  
P.O. Box 80.021  
3508 TA Utrecht  
The Netherlands**

*Voor mijn ouders*

*Contents*

**Contents**

**Chapter 1**  
*Introduction and summary* ..... 1

**Chapter 2**  
*Are long-period body-wave coda caused by lateral heterogeneity? - I. the coda of P* ..... 7

    2.1 Introduction ..... 8

    2.2 Data ..... 8

    2.3 Spatial coherency of the signal ..... 10

    2.4 Coda predicted by spherically symmetric models ..... 12

    2.5 Upper mantle discontinuities ..... 19

    2.6 Sensitivity to LVZ structure ..... 20

    2.7 Discussion ..... 21

    References ..... 23

**Chapter 3**  
*Are long-period body-wave coda caused by lateral heterogeneity? - II. the coda of PP* ..... 27

    3.1 Introduction ..... 27

    3.2 Data ..... 29

    3.3 Slowness and azimuth analysis ..... 30

    3.4 Frequency content of coda waves ..... 35

    3.5 Body wave scattering ..... 36

    3.6 Rayleigh to P coupling ..... 38

    3.7 Discussion ..... 43

    References ..... 44

    Appendix A: beamforming ..... 46

## Contents

### Chapter 4

<i>Topography of the 400 km discontinuity from observations of long-period <math>P_{400}</math> <math>P</math> phases</i> .....	51
4.1 Introduction .....	52
4.2 Data .....	52
4.3 Data processing .....	58
4.4 Possible sources of error .....	61
4.5 Focusing of $P_d$ $P$ phases .....	62
4.6 Models of topography .....	63
4.7 Conclusions .....	73
References .....	75

### Chapter 5

<i>A Formalism for Including Amplitude Data in Tomographic Inversions</i> .....	77
5.1 Introduction .....	78
5.2 Perturbation theory for amplitudes .....	79
5.3 Validity of the method .....	90
5.4 Application to the upper mantle .....	91
5.5 Conclusions .....	97
References .....	98
Appendix A. Explicit form of equation (10). .....	100
Appendix B. A different approach to find the amplitude perturbation. ....	100

### Chapter 6

<i>The use of P-wave amplitude data in a joint tomographic inversion with travel times for upper mantle velocity structure</i> .....	103
6.1 Introduction .....	104
6.2 Data selection and processing .....	106
6.3 Joint inversion of travel-time and amplitude data .....	114
6.4 Joint inversion of synthetic travel-time and amplitude data .....	119
6.5 Amplitudes as a validation tool .....	128
6.6 Discussion .....	131
References .....	133
Appendix. The effect of attenuation .....	134



## *Contents*

<i>Samenvatting</i> .....	137
<i>Acknowledgements</i> .....	141
<i>Curriculum Vitae</i> .....	143

## Chapter 1

### Introduction and summary

The last decade has seen a large increase in the amount of high-quality data from a growing number of digitally operating seismic stations. Both short- and long-period data are reported on a regular basis to central data centres. As global seismology is hampered by the unequal distribution of stations and earthquakes, extracting as much as possible information from the available data is important. However, only a relatively small part of the growing data volume is used on a more or less routine basis. New sources of information from the seismogram may add independent constraints on and provide a higher resolution of Earth structure.

The most widely used parts of the seismogram are the  $P$ -wave travel time and the long-period surface wave train. The travel time of the (mostly short-period)  $P$ -wave is used on a large scale to invert for Earth structure (Spakman 1991; Vandecar 1991; Van der Hilst & Engdahl 1991). Recently, also the  $S$ -wave travel time is used in tomographic inversions (e.g., Zielhuis 1988). On the low-frequency end of the spectrum, long-period surface wave forms are inverted for mantle velocity structure (e.g., Woodhouse & Dziewonski 1984; Snieder 1988). Recent studies include the long-period  $S$ -waves in the wave form modeling using modal summation (Nolet 1990; Zielhuis 1992). The results of Li & Tanimoto (1991) show that it is still prohibitively time-consuming to include also the long-period  $P$ -wave in this process.

This leaves large parts of the seismogram unused in inversion procedures. Before these can be included in any inversion process, their information content and relation to Earth structure must be established (e.g., wave propagation in a laterally homogeneous or heterogeneous Earth model). The main goal of the research presented in this thesis is to determine the information content of parts of the seismogram that are not frequently used and, if possible, to develop methods to extract this information.

In the first part of this thesis, special emphasis is put on the intervals between the different body wave arrivals (the body-wave coda) in the intermediate and long-period seismogram. The periods contained in the data are longer than about 7 s. The wave forms of these body waves ( $P$ ,  $PP$ ,  $S$ , etc.) are used to construct laterally homogeneous models of upper mantle layering (e.g., Given & Helmberger 1980; Burdick 1981; Paulssen 1988). The intervals between the body waves make up a considerable portion of the seismogram and are a

potentially rich source of information. At long periods, these body-wave coda intervals may be expected to be less affected by scattering processes than short-period seismograms, and may therefore be amenable to, e.g., waveform modeling. Numerous studies have shown the dominance of scattering processes in short-period (around 1 s) seismograms (e.g., Aki & Chouet 1975; Sato 1984), which makes it necessary to use statistical methods to extract information on Earth structure from the short-period component of the data.

An analysis of data from the Network of Autonomously Recording Seismographs (NARS) and Graefenberg (GRF) arrays in western Europe shows a marked difference in behaviour between the coda following the  $P$ -wave (that part of the seismogram between the  $P$  and  $PP$ -wave) and the coda following the later arriving body waves. Whereas the  $P$ -wave coda is coherent between stations of a small array (GRF) and shows no signs of wave scattering, the later body-wave coda intervals are dominated by incoherent energy. A newly developed beamforming analysis is used to analyse these incoherent signals. This yields information not only on the wave types present, but also on the direction of propagation of these waves. Wave scattering and/or conversion is reflected in deviations from great circle azimuth. The incoherent behaviour of the  $PP$  and  $S$ -waves coda is shown to be due to body wave to surface wave scattering. Some simple calculations show that the most likely location of the scatterers is near or at the surface (surface topography). The coherent behaviour of the  $P$ -wave coda is used to extract constraints on a laterally homogeneous upper mantle model for the great circle path between source and station. The results suggest that it may be possible to use the  $P$ -wave coda wave forms to study, for example, variations in the low-velocity zone.

The results of chapters 2 and 3 show that the signal in the intermediate and long-period  $P$ -wave coda can be explained by wave propagation in a laterally homogeneous Earth model, whereas the later coda intervals are dominated by scattered energy, requiring laterally heterogeneous Earth models. Therefore, the  $P$ -wave coda may be used to invert for upper mantle structure between source and station, whereas the information content and probably complex relation with Earth structure of the later coda intervals make an inversion of these parts unattractive.

In chapter 4 an example of the analysis of the  $P$ -coda is presented. A reflection at the underside of the 400-km discontinuity halfway between source and station ( $P_{400}P$ ) is used to study this discontinuity. This phase arrives in the coda of  $P$ . Observed variations in the amplitude of  $P_{400}P$  phases that are reflected near Hawaii are interpreted as the result of focusing by an undulating 400-km discontinuity. A Kirchhoff integral representation is used to describe the focusing. As the  $P_{400}P$  phases are sensitive to a large area of the discontinuity, it is not possible to obtain a unique, deterministic model of the topography of the discontinuity with the number of data available. Instead, a range of wave lengths and amplitudes of topography that can explain the amplitude fluctuation in the data is given. Topography with wave lengths of about 500-1500 km and amplitudes of 15-20 km are consistent with the data. No signature of a possible plume of hot mantle material beneath the Hawaii hotspot region could be found in the data; it is shown that the amplitude signal of such a structure is much smaller than observed amplitude variations. If more  $P_{400}P$  data are available, these could be used to study the upper mantle discontinuity for a signature of upper mantle convection in the topography of the mantle discontinuities, and possibly constrain convection cell size.

The results from chapter 4 show that although it is possible to explain the character of the

*P*-coda with laterally homogeneous Earth models (chapter 2), a laterally varying layering is necessary to describe the behaviour of  $P_{400}P$ . If enough data are available, it may be possible to map these variations. An interesting result is that for these reflections at the underside of an upper mantle discontinuity, it is possible to explain variations in the amplitude of the  $P_{400}P$  phase on a scale smaller than the size of the Fresnel zone. This surprising result can be explained by the small curvature of the  $P_{400}P$  travel-time surface near the reflection point. Relatively smooth topography can induce multiple points of stationary travel time, resulting in multipathing and, possibly, small-scale amplitude fluctuations.

The second part of the thesis deals with the amplitude of short-period *P*-waves. Although this quantity is relatively easy to obtain, it is not used often, due to both the number and complexity of the processes that affect the amplitude of a *P*-wave. Furthermore, a large data set is needed to reduce the large scatter observed in *P*-wave amplitude data.

A number of attempts to invert *P*-wave amplitude data have been reported. In attenuation tomography (see, e.g., Ho-Liu et al. 1988), observed *P*-wave amplitudes variations are attributed entirely to variations in the *Q*-factor. Geometrical spreading is ignored, which may be a questionable assumption for the mostly complex media to which these studies are applied. Thomson (1984) interpreted short-period *P*-wave amplitudes in terms of geometrical spreading in a complex medium. The results were unconvincing, due to both moderate data quality and probably the break-down of the linearization in the theory relating slowness and amplitude perturbations.

In chapters 5 and 6 the use of *P*-wave amplitudes in inverse modeling is studied, interpreting the amplitudes in terms of geometrical spreading through a three-dimensionally varying medium. Chapter 5 presents an efficient (linearized) method to calculate Frechet derivatives of amplitudes for slowness perturbations. With this theory, it is possible to handle large amplitude data sets in tomographic inversions. Calculations with synthetic data using a realistic model of upper mantle slowness variations show the strongly non-linear relation between amplitude perturbations and slowness perturbations along the ray path, restricting the applicability of the linearized theory. This depends on the geometry of the ray and the scale length of the slowness variations.

The results of joint inversions of observed and synthetic travel-time and amplitude data using the theory developed in chapter 5 are presented in chapter 6. The processing of *P*-wave amplitudes is shown in detail, as this is a key element in obtaining a high-quality data set. As amplitude data show large scatter, it is important to remove outliers. These may seriously influence the result of a least-squares inversion. It is shown that the *P*-wave amplitudes are dominated by effects originating near the station. Correcting for these effects is important, as the effects of wave propagation through the mantle are masked by these local effects.

The results of the joint inversions show that adding amplitude data to travel-time data does not increase the resolution, even though tests with synthetic data show that the travel times alone have difficulty in resolving sharp gradients in some parts of the synthetic model. The amplitude data induce small, short-scale slowness fluctuations in the model. These increase the roughness of the model only slightly, and are therefore not affected by the smoothness constraint used to regularize the inversion. Two reasons are put forward to explain this result. Firstly, in the setup used in chapter 6, with teleseismic *P*-waves and relatively small-scale (about 30 km) slowness variations allowed in the model, the geometrical spreading amplitudes appear to be extremely sensitive to slowness perturbations in the

upper mantle below the station. This has the effect that slight, small-scale perturbations throughout the model are sufficient to match the amplitude data. Secondly, the coverage of the model by the data set used is apparently not complete enough to induce longer-wavelength changes to the model.

An additional disadvantage of amplitude data is the large variance of these data. To decrease the variance in the amplitude data to the level of the variance in the travel times, the amplitude data set should be at least an order of magnitude larger than that used in chapter 6. This would prevent downweighting of the amplitude data in inversions based on a Bayesian formulation. Therefore, it is concluded that body-wave amplitude data are useful in upper mantle tomographic inversions with travel times, only if a large number of high-quality data is available to ensure both a good coverage of the model and low variance.

Instead of using the data in inverse modeling, the use of amplitudes as a validation tool is suggested, for example as a check on models obtained with different data set(s). Body-wave amplitude data may be more useful in different setups, where the ratio of ray length and scale length of slowness variations affecting the amplitudes is smaller. In that case the sensitivity of the amplitudes to small velocity changes along the ray path is smaller. This may be the case in cross-borehole experiments, or in studies of long-period surface waves.

## References

- Aki, K. & B. Chouet, 1975. Origin of coda waves: source, attenuation and scattering effects, *J. Geophys. Res.*, **78**, 3322–3342.
- Burdick, L.J., 1981. A comparison of the upper mantle structure beneath North America and Europe, *J. Geophys. Res.*, **86**, 5926–5936.
- Given, J.W. & D.V. Helmberger, 1980. Upper mantle structure of northwestern Eurasia, *J. Geophys. Res.*, **85**, 7183–7194.
- Ho-Liu, P., H. Kanamori & R.W. Clayton, 1988. Applications of attenuation tomography to Imperial Valley and Coso-Indian Wells region, southern California, *J. Geophys. Res.*, **93**, 10,501–10,520.
- Li, A.A. & T. Tanimoto, 1991. Waveforms of long-period body waves in a slightly aspherical Earth model, *Geophys. J. Int.*, *submitted*.
- Nolet, G., 1990. Partitioned waveform inversion and two-dimensional structure under the Network of Autonomously Recording Seismographs, *J. Geophys. Res.*, **95**, 8499–8512.
- Paulssen, H., 1988. Lateral heterogeneity of Europe's upper mantle as inferred from modelling of broad-band body waves, *Geophys. J. R. astron. Soc.*, **91**, 171–199.
- Sato, H., 1984. Attenuation and envelope formation of three-component seismograms of small local earthquakes in randomly inhomogeneous lithosphere, *J. Geophys. Res.*, **89**, 1221–1241.
- Snieder, R., 1988. Large scale waveform inversions of surface waves for lateral heterogeneity - II: Application to surface waves in Europe and the Mediterranean, *J. Geophys. Res.*, **93**, 12067–12080.
- Spakman, W., 1991. Delay time tomography of the upper mantle below Europe, the Mediterranean, and Asia Minor, *Geoph. J. Int.*, **107**, 309–332.
- Thomson, C., 1983. Ray-theoretical amplitude inversion for laterally varying velocity structure below NORSAR, *Geophys. J. R. astr. Soc.*, **74**, 525–558.
- Thomson, C.J. & D. Gubbins, 1982. Three-dimensional lithospheric modelling at NORSAR: linearity of the method and amplitude variations from the anomalies, *Geophys. J. R. astr. Soc.*, **71**, 1–36.
- Vandecar, J.C., 1991. *Upper mantle structure of the Cascadia subduction zone from non-linear teleseismic travel-time inversion*, PhD thesis, University of Washington, Seattle, Washington.
- Van der Hilst, R.D. & E.R. Engdahl, 1991. On ISC PP and pP data and their use in delay-time tomography of the Caribbean region, *Geophys. J. Int.*, **106**, 169–188.

- Woodhouse, J.H & A.M. Dziewonski, 1984. Mapping the upper mantle: three-dimensional modeling of Earth structure by inversion of seismic waveforms, *J. Geophys. Res.*, **89**, 5953–5986.
- Zielhuis, A., 1988. A shear velocity model for Europe from ISC delay time data, in *Proceedings of the Fourth EGT Workshop*, edited by Nolet, G. and B. Dost, pp. 59-62, European Science Foundation, Strasbourg.
- Zielhuis, A., 1992. *S-wave velocity below Europe from delay-time and waveform inversions*, PhD thesis, University of Utrecht, Utrecht.

## Chapter 2

### Are long-period body-wave coda caused by lateral heterogeneity? - I. The coda of *P*

**Summary.** Data from two broad-band arrays in western Europe (NARS and GRF) are used to study the character of long-period coda of *P*. The events studied are at epicentral distances of 40° to 60°, in the Hindu Kush region and on the Mid-Atlantic Ridge, sampling the upper mantle to a depth of about 1000 km. The periods studied are 5-50 s. In contrast to observations at much shorter periods (1 s) the long-period coda of *P* at the GRF array (inter-station distance about 10 km) are strongly coherent. Synthetic seismograms obtained with the reflectivity method show that spherically symmetric upper mantle models can explain this part of the data; it is not necessary to invoke scattering of any kind. An upper mantle with a thick lid overlying a pronounced low-velocity zone (LVZ) is necessary to explain the character of the *P* coda at GRF. Such an upper mantle model agrees with previous studies of similar great circle paths. The strong coherency of the *P* coda is lost on the scale of NARS (station separation about 200 km); this indicates lateral variations in the upper mantle at a scale of about 200 km. It appears from previous studies of the upper mantle under Europe that these variations must be sought in the LVZ. It is shown that the long-period *P* coda is sensitive to variations in the *P*-velocity structure of the LVZ. This suggests the *P* coda as a tool for monitoring lateral variations in the LVZ and possibly to prove the existence or absence of a LVZ in the *P*-velocity.

---

Chapters 2 and 3 have been published as

Neele, F. and R. Snieder, Are long-period body wave coda caused by lateral inhomogeneity?, *Geophys. J. Int.*, **107**, 131-153, 1991.

## 1. Introduction.

In chapters 2 and 3 the long-period body wave coda are used to infer which wave propagation phenomena are responsible for different parts of the seismogram. The central problem is to find out what class of Earth models is consistent with the data and whether different parts of the seismogram require different classes of models (spherically symmetric versus laterally heterogeneous). The aim is not to find models that exactly fit the data, but rather to find constraints on possible models.

The body wave coda have received much attention in the literature. Both laterally heterogeneous and homogeneous upper mantle models have been used to explain the body wave coda. In a spherically symmetric model, top-side reverberations on the upper mantle discontinuities, either single or multiple, cause coda waves, such as the leaky, long-period *PL* mode (Oliver & Major 1960; Helmberger & Engen 1980; Baag & Langston 1985) or single reverberations observed in short-period *P* coda (e.g., Husebye & Madariaga 1970; Gutowski & Kanasewich 1974); bottom-side reflections produce precursors to *PP* (Bolt 1970; Wajeman 1988; Shearer 1990) or *P'P'* (Whitcomb & Anderson 1970; Husebye et al. 1977; Nakanishi 1988; Davis et al. 1989). Waves may also be converted at upper mantle discontinuities (e.g., Barley et al. 1982; Paulssen 1985, 1988b; Kind & Vinnik 1988). For short periods (shorter than about 1 s) scattering has been successfully used to explain observed *P*-wave coda characteristics, such as duration and amplitude (e.g., Aki 1973, 1980ab; Frankel & Wennerberg 1987; Kennett 1987; Flate & Wu 1988; Korn 1988); the scattering approach has also been applied to the short-period *P* coda to explain it as precursors to *PP*, originated from scattering of *PP*-waves in the upper few hundred kilometers of the upper mantle (King et al. 1975).

With the exception of those of the *PL* mode, all studies cited above considered short periods. The dominant mechanism to create short-period body wave coda appears to be scattering at lateral heterogeneity or reflection and/or conversion at discontinuities in the upper mantle. At long periods reverberations in a laterally homogeneous Earth (the *PL* mode) offer an adequate explanation of that part of the coda immediately following the direct *P*-wave.

Recently, digital broad-band data have become available from networks in Europe, with a large variety of station separations. These data are used to study coda generation. In this chapter the coda of *P* (which includes precursors to *PP*) is analyzed. The coda of *PP* (the interval between the *PP*-wave train and the *S*-wave) is the subject of chapter 3. Epicentral distances range from 40° to about 60°, so essentially the lithosphere and upper mantle structure to a depth of about 1000 km is probed. Synthetic seismograms are used to show that laterally homogeneous models of the upper mantle can explain the long-period *P* coda and to put constraints on upper mantle structure.

## 2. Data.

The data used in this study are from the digital, broad-band GRF (Gräfenberg, West Germany) and NARS arrays. The locations of the events are shown in figure 1; table I lists the event parameters. We are interested in long-period body waves and consider periods in the range of about 5-50 s.

The GRF array is L-shaped, about 100 km long in the N-S direction, while the E-W extent is about 70 km (figure 2). The array consists of 13 broad-band seismometers, of



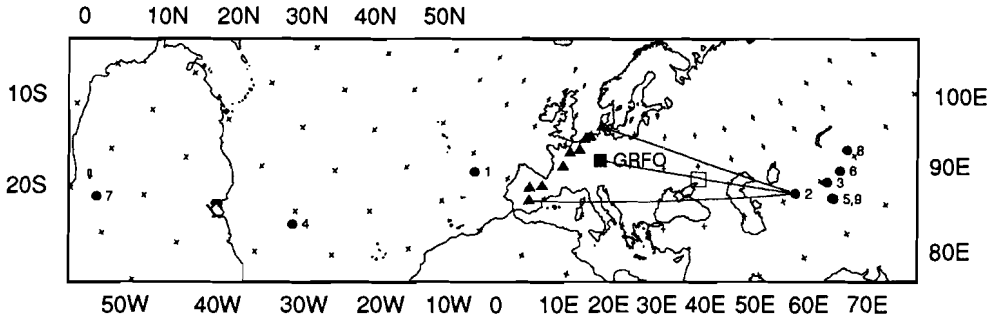


Figure 1. The events used in this study (solid circles). Also shown is the position of the Graefenberg array (solid square) and the NARS array, of which a number of stations are plotted (solid triangles). The great circle of event 2 from table I to GRF is plotted, together with the great circles to the outermost NARS stations. The open square represents the bounce point of  $P_dP$  phases to GRF from event 2.

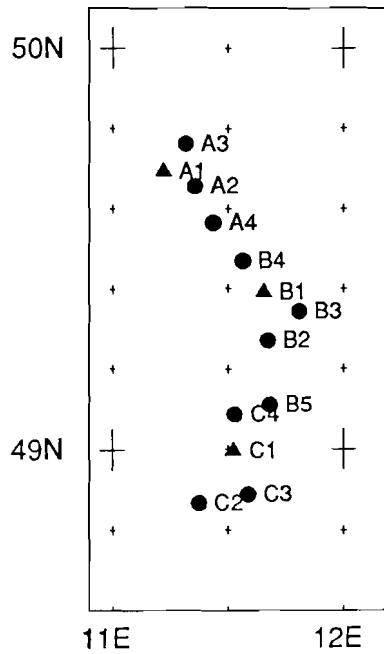


Figure 2. The Graefenberg array. Solid circles are 1-component stations, solid triangles denote 3-component stations.

which 3 have 3 components. The station separation is typically about 10 km. The NARS array extends from the southernmost part of Sweden to the south of Spain, with an average station separation of 200 km (figure 1). This allows a comparison of the characteristics of the wave field at two scale lengths.

Table I. Event information.

#	Date	Time	Lat	Lon	h	$m_b$	Dist
1	1983-10-17	19:36:21.6	37.590	-17.410	10	5.9	23.8°
2	1984-03-19	20:28:38.3	40.350	63.360	15	6.4	37.1°
3	1984-10-26	20:22:18.0	39.150	71.350	7	5.9	42.9°
4	1984-11-01	04:48:50.5	8.160	-38.770	10	6.3	58.7°
5	1985-07-29	07:54:44.5	36.210	70.900	102	6.5	44.3°
6	1985-08-23	12:41:56.4	39.490	75.270	40	6.2	45.2°
7	1986-05-09	16:23:48.8	-17.170	-65.620	13	5.6	94.9°
8	1987-01-24	08:09:21.8	41.518	79.293	33	6.0	46.6°
9	1987-05-05	15:40:48.2	36.464	70.683	211	5.7	44.0°

### 3. Spatial coherency of the signal.

The coherency of the seismic wave field, both as a function of frequency and station separation, contains information on the scale lengths of inhomogeneities in the upper mantle. Many studies have been devoted to the coherency of short-period body waves at seismic arrays (e.g., Aki 1973; King et al 1975; Kennett 1987; Korn 1988; Toksöz et al 1989) to obtain information about the structure under the array.

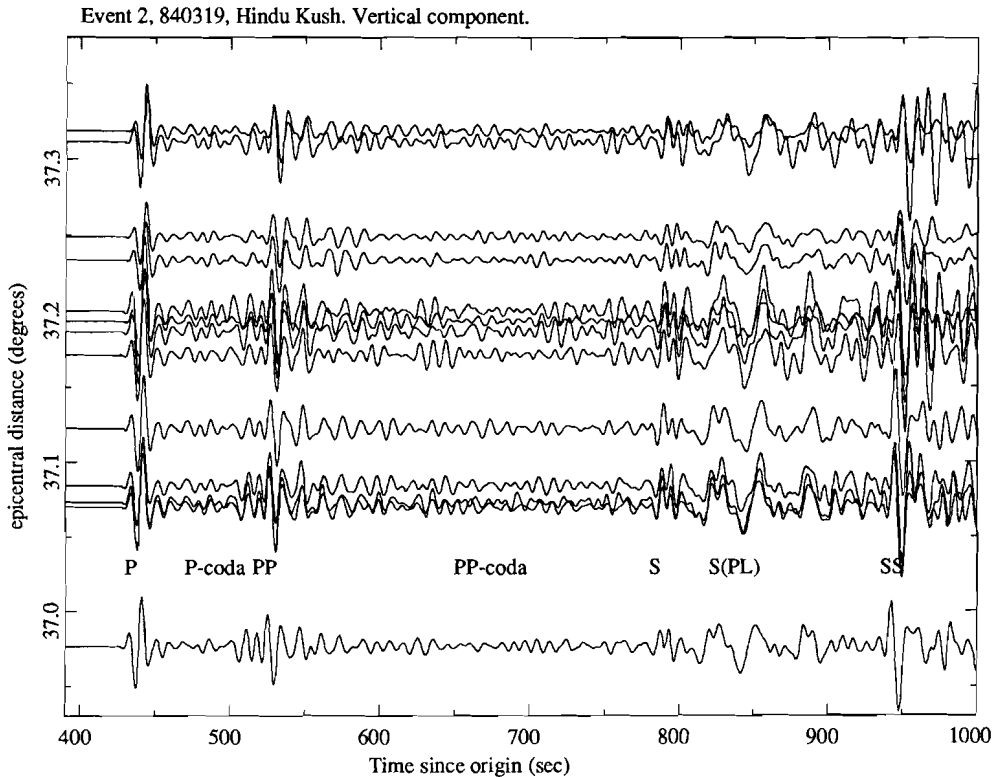
Figure 3 shows a time-distance plot of event 2 recorded by GRF. The traces have been lowpassed with corner frequency of 0.15 Hz. The recordings for this event are representative for all events in table I. The coherency of the traces up to about 50 s after the *PP*-waves is clear. Almost every wave form can be traced from one station to the next. Note that the noise level is low, therefore all signal is related to the event and therefore due to the response of the Earth.

In contrast, the *PP* coda, between the *PP*-waves and the *S*-wave, is clearly dominated by incoherent energy, which is evidence for strong scattering. At this distance a long-period and large-amplitude *S*-coupled *PL*-wave constitutes the *S* coda, masking incoherent arrivals. Assuming that scattering in the lower mantle does not contribute significantly to the wave scattering, the incoherent energy must be caused by scattering somewhere in the upper mantle near the receiving array; scattering near the source will result in coherent energy at this distance and can only be separated from the primary wave field by detailed slowness and azimuth analysis. This has been done for the short-period *P* coda at 100° by King et al. (1975).

Figure 4 shows the data for the same event, recorded by the NARS array. The data are lowpassed at 0.15 Hz. The coherency of the *P* coda is no longer present at these station separations, the only coherent arrivals seem to be the main body wave phases. The first part of the *PP* coda that is coherent at GRF is incoherent between NARS stations. This incoherence is not surprising, as the slowness of *PP* (and thus the excitation of *PP*-coupled *PL* (Alsop & Chander 1968)) varies across such a large array; furthermore the *PP*-waves are in the triplication zone, causing a rapidly changing *PP*-wave train. However, note the data from stations NE03 and NE04 at the bottom of figure 4: these stations have the same epicentral distance, but they show large differences in the *PP*-wave train and, to a smaller extent, in the *P* coda. Note also stations NE13 and NE14 at the top of figure 4. These are close in epicentral distance, but the *P* coda is incoherent between the stations.

The *P* coda of event 4 on GRF records is shown in figure 5. The noise level is low and there is hardly any incoherency in these data. A *PcP* phase arrives at  $t=650$  s.

The coherency of the *P* coda on the GRF records implies that the importance of



*Figure 3.* Time-distance plot of event 2 recorded by GRF. The arrival time of major phases is indicated. The data are low-passed at 0.15 Hz. Note the strongly coherent *P* coda, and the incoherent coda of *PP*.

scattering is small. A slowness-azimuth analysis using a beamforming algorithm especially suited for long-period body wave analysis showed no detectable deviations from great circle azimuth in the *P* coda (see chapter 3 for a detailed description of the algorithm and results). This proves that scattered waves, if any, are weak. The data in this part of the seismogram should be reproducible by a spherically symmetric Earth model. The observation that the *P* coda is coherent at a scale of hundred kilometers and that this coherency is lost at a scale of more than a few hundred kilometers, suggests that there exist lateral variations in the mantle on a scale comparable to the station separation of the NARS array. At the epicentral distances of the events in table I, the waves penetrate the mantle to a depth of about 1000 km, so these large-scale lateral variations must be situated in the upper mantle, assuming a largely homogeneous lower mantle. A recent stochastic analysis of global travel time data supports this assumption: a heterogeneous upper mantle and a low level of heterogeneity in the lower mantle (Gudmundsson et al. 1990). Figure 1 also shows some great circle paths from event 2 to GRF and some NARS stations. The great circle paths to the NARS stations are so far apart that they may sample completely different tectonic regions. Thus, incoherency of the *P* coda is expected. Tomographic studies in Europe using *P*-wave delay times (Spakman 1991) or waveforms of the fundamental Rayleigh mode (Snieder 1988)

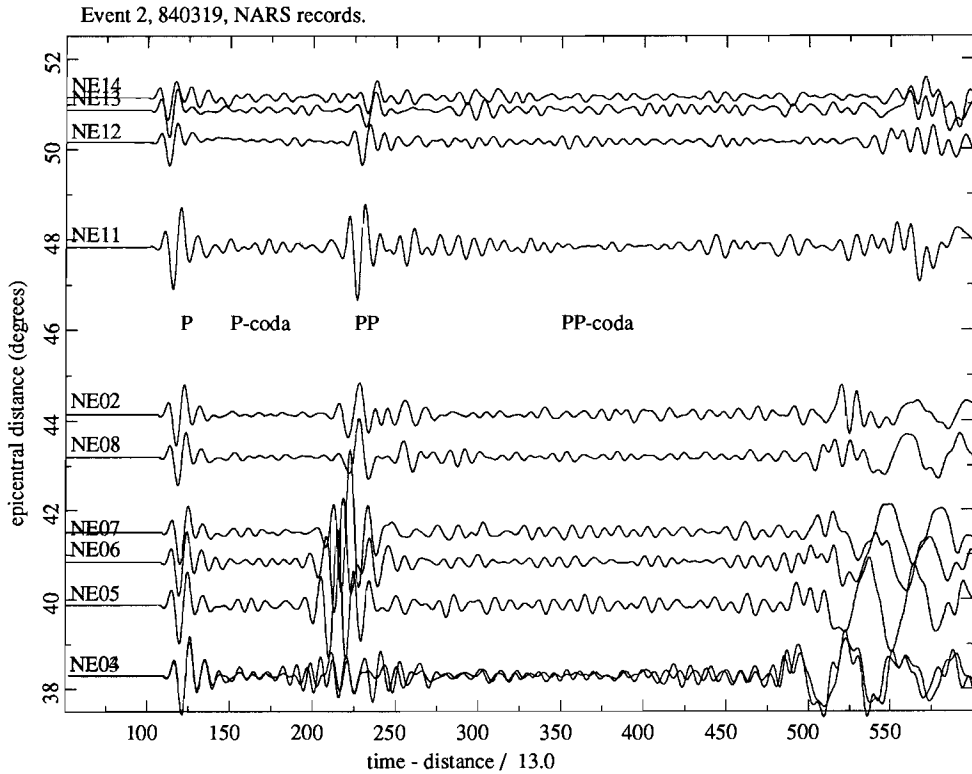


Figure 4. Time-distance plot of the event from figure 3, now recorded by the NARS array. The data are low-passed at 0.15 Hz. The coherence of the *P* coda that is present in the GRF data is absent at this array.

also show large lateral variations in the upper mantle on a scale of several hundred kilometers, comparable to the NARS station separation. In the next section the *P* coda is used to constrain the upper mantle between source and station.

#### 4. Coda predicted by spherically symmetric models.

The strong coherency of the *P* coda suggests that a laterally homogeneous Earth model can be found that satisfies the data. A number of laterally homogeneous upper mantle models was taken from the literature to compute synthetic seismograms, to find out to what extent these models can explain the data. Only data from three shallow events (events 2, 3 and 4 in table I) are used, so that the *pP* and *sP* phases do not interfere with the *P* coda. In this case of only one source-receiver path for each event, the problem of non-uniqueness is large; therefore, it is not attempted to fit the data perfectly. A reflectivity code of Kennett (1988) is used, which ensures that the total response is obtained.

In order to quantify the differences between synthetics and data, two simple measures of this difference are calculated. One is the energy in the coda of *P*. As can be seen in figures 3 and 5, the data show considerable energy in this interval. Ignoring scattering near the

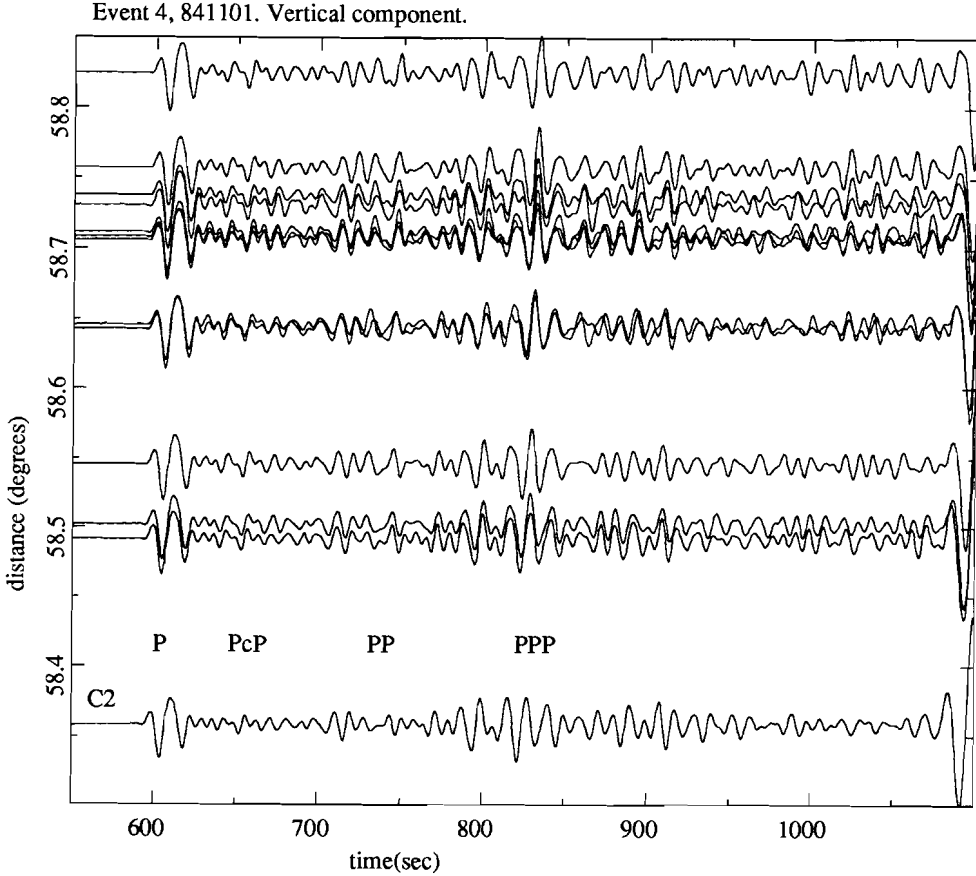


Figure 5. Time-distance plot of event 4, recorded by GRF. This event shows a large coda level. Again the *P* coda is coherent, whereas the *PP* coda is largely incoherent across the array. The *S*-wave arrives just at the end of the section.

source, the upper mantle model is constrained to generate this amount of energy. As pointed out in the introduction, the signal from the 1-D earth between *P* and *PP* comprises top-side reverberations (with slightly higher slowness than that of *P*) and bottom-side reflections (with slowness slightly lower than that of *PP*) on the upper mantle discontinuities. Therefore, the energy in the *P* coda is normalized to the energy in both the *P* and *PP*-waves:

$$f = \frac{|P \text{ codal}|^2}{|P \text{ wave}|^2 + |PP \text{ wave}|^2} \quad (1)$$

where  $| \cdot |^2$  denotes the energy in the interval between brackets. The value  $f$  can be thought of as a measure of the efficiency of a model to transfer energy of the *P* and *PP* phases into the *P* coda. As it is not attempted to model details in the data, the relative timing of

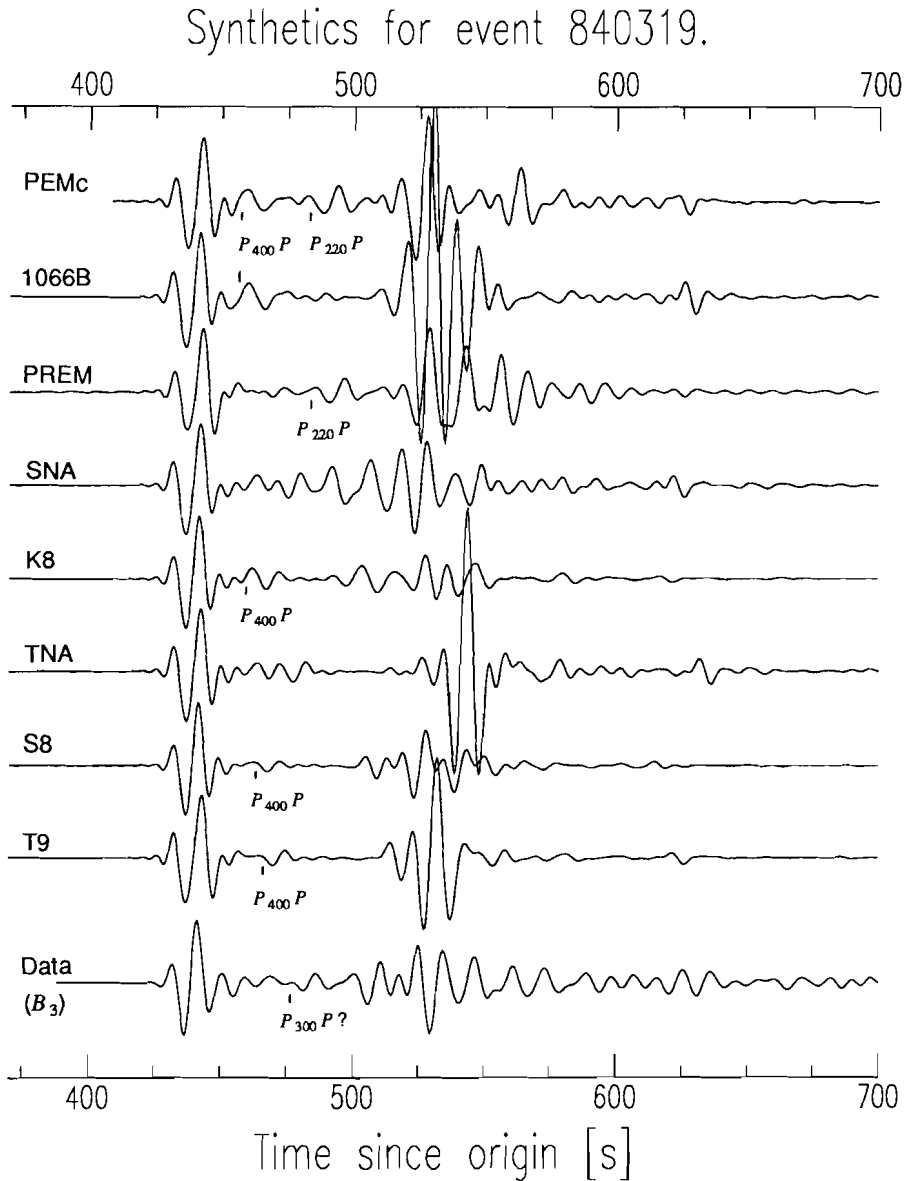
individual phases in the coda of  $P$  is of no interest. The  $P$  coda measure described here is independent of such timing differences.

Another quantitative measure of the difference between synthetics and data is the energy in the coherent part of the  $PP$  coda. Figure 3 shows that the observed  $PP$  coda has reverberative characteristics and is reasonably coherent across the array up to about 600 s. The energy in the coherent part the  $PP$  coda interval is normalized by the energy in the  $PP$  interval.

Reflectivity synthetics are computed for the following models: global models 1066B (Gilbert & Dziewonski 1975), PREM (Dziewonski & Anderson 1981) and PEMc (Dziewonski et al. 1975) from normal mode data and body-wave travel times; a number of local models derived from body wave form analysis: models S8 (Burdick 1981) and K8 (Given & Helmberger 1980) for western and eastern Europe, respectively; models SNA and TNA (Grand & Helmberger 1984), T9 (Burdick 1981) for the stable and tectonic parts of North America. If for these models only one of  $P$  and  $S$ -velocity is given, the other is calculated using the  $P/S$  ratio of PREM. All models, except the global models, are extended below the upper mantle with the lower mantle of PREM. To include attenuation, the  $Q$ -model given by Der et al. (1986) for shield areas under Eurasia derived from body wave data is used. In the reflectivity computations, the layer thickness in the models is taken to be no more than 0.5 times the shortest wavelength at each depth (see Chapman & Orcutt 1985 and references therein for a discussion on this point). The sampling of the slowness integral in the reflectivity calculations is always more than 10 points per shortest wavelength (Mallick & Frazer 1987). Both slowness and frequency integrals are tapered to avoid cut-off artifacts. The synthetics are then lowpassed at 0.15 Hz. The source mechanism is taken from the ISC bulletins. The GRF instrument response was applied to the synthetics, which were then compared with the data. After a convolution with the GRF instrument response, the synthetics can be compared with the data.

Figures 6, 7 and 8 show a series of synthetic seismograms for each of the three events. The synthetics have been shifted in time to line up with the first onset in the data (bottom traces in figures 6, 7 and 8, which represent the data from one of the stations of GRF). Table II lists the measures of  $P$  and  $PP$  coda energy.  $P$  and  $PP$  coda energy measures are obtained for the data as described in equation (1); for each event an average over all stations of GRF is obtained. The energy measures for the synthetics are normalized by these observed average energy measures. The data thus have a value of 1.0; a value lower than 1.0 means that the energy in the synthetics is lower than what is present in the data. The results for event 2 (figure 6) show that a number of models generates a  $P$  coda that is comparable to the data. These models are K8, PEMc and PREM. Model SNA predicts a coda that is too large. The other models do not fit the data well in this respect. Comparing the models, it appears that the models that best fit the data all have a thick lid (about 100 km) and a more or less pronounced low velocity zone (LVZ) in both  $P$ - and  $S$ -velocity. In these models, the jump in velocity at the base of the lid is about 2-4%. Models with no LVZ or a very thin lid (as TNA, S8, T9) do not generate a sufficiently strong  $P$  coda.

Figure 7 and table II show the results for event 3, which is close to event 2. The difference between the synthetics in figures 6 and 7 is entirely due to the difference in epicentral distance. Synthetics calculated with an explosion source show comparable differences. As again models with a thick lid and a LVZ (SNA, PEMc; 1066B has a LVZ in the  $S$ -velocity) show the best fit, the average upper mantle between GRF and the two events is essentially



**Figure 6.** Reflectivity synthetics for a number of spherically symmetric upper mantle models. Bottom trace represents data from event 2 at station B3 of GRF. The synthetics have been shifted to line up the P-waves with the first onset in the data. Synthetics and data are low-passed at 0.15 Hz. The traces are normalized to their maximum amplitude; some phases are indicated.

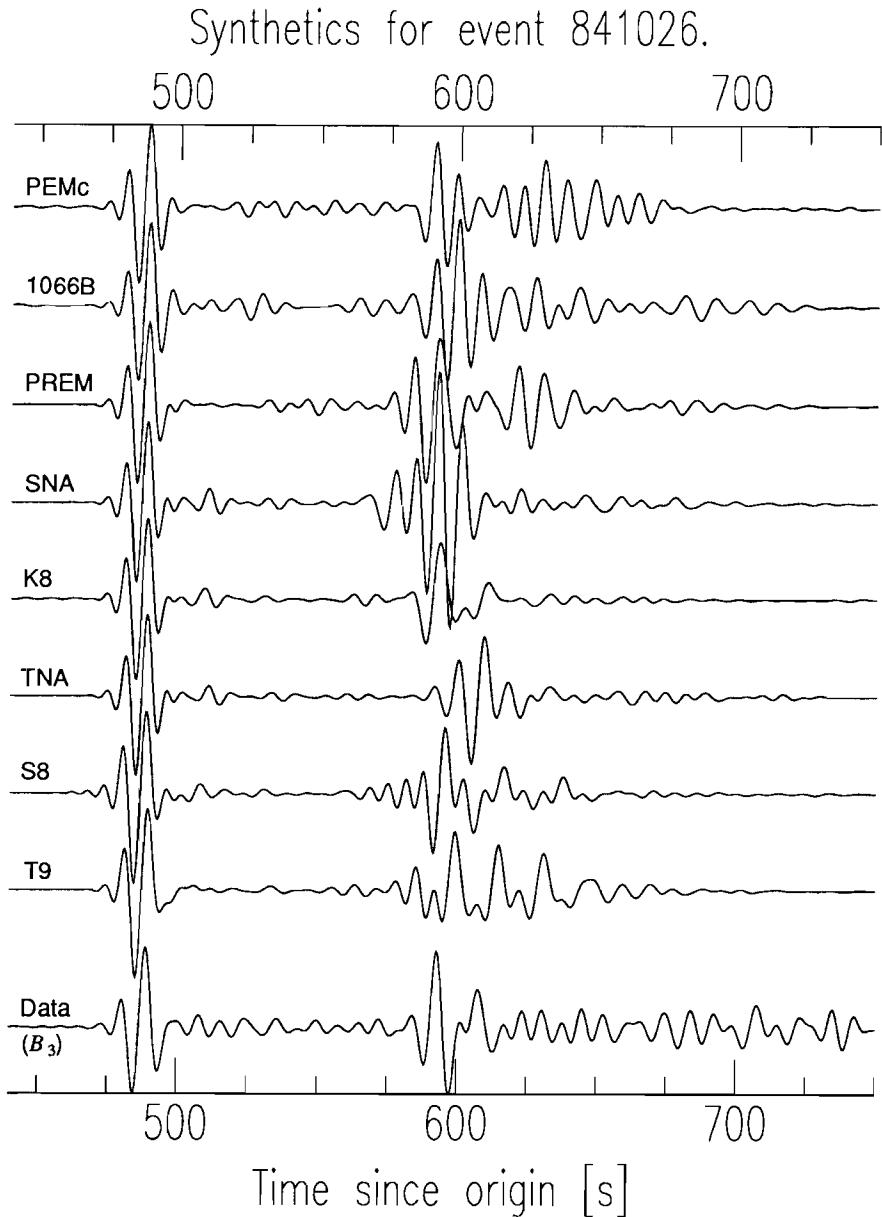


Figure 7. As figure 6, now for data from event 3 (station B3).

the same. The great circle paths are both within the stable shield areas in Europe (see Rial et al. 1984).

The second quantitative measure is the length and amplitude of the coherent part of the *PP* coda. From figure 6 and 7 it is apparent that the *PP* coda strongly depends on the model. From visual inspection and table II it follows that PEMc seems the most



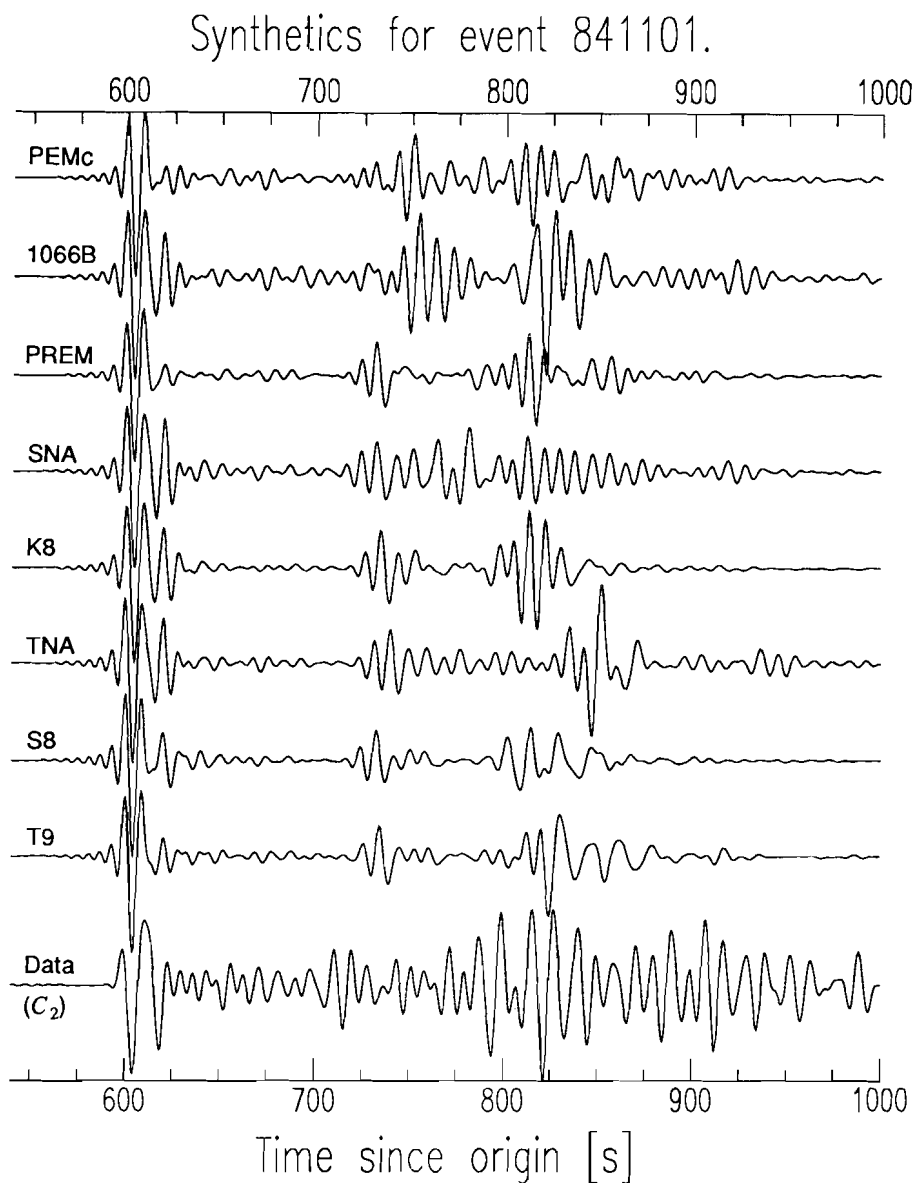


Figure 8. As figure 6, now for data from event 4 (station C2).

appropriate model for the upper mantle between GRF and Hindu Kush. It should be noted here, that if a smaller  $P/S$  ratio is assumed for SNA, both the  $P$  and  $PP$  coda generated by SNA are much closer to the data. The models that fit the observed  $PP$  coda best are again characterized by a thick lid over a LVZ, with a more or less pronounced jump at the base of the lid.

Table II. Measure of energy in  $P$  and  $PP$  coda.

Model	Event 2		Event 3		Event 4	
	$P$ coda	$PP$ coda	$P$ coda	$PP$ coda	$P$ coda	$PP$ coda
PEMc	0.82	0.79	0.43	0.90	0.13	0.27
1066B	0.13	0.02	0.34	0.32	0.27	0.17
PREM	0.57	1.66	0.29	0.25	0.05	0.08
SNA	2.72	0.33	0.41	0.14	0.46	0.19
K8	0.78	0.19	0.27	0.09	0.16	0.05
TNA	0.20	0.05	0.32	0.18	0.33	0.17
S8	0.25	0.25	0.23	0.24	0.16	0.19
T9	0.19	0.04	0.16	0.42	0.13	0.27
Data	1.00	1.00	1.00	1.00	1.00	1.00

Event 4 lies on the Mid-Atlantic Ridge (MAR). Synthetics and data are shown in figure 8. The data contain a  $PcP$  arrival at 650 s, which is not included in the synthetics. This will bias the results for the  $P$  coda fit to low values. Although this path is for the larger part oceanic, again a model with a thick lid and LVZ is found that explains the data best. The average upper mantle here would not be expected to be similar to GCA (Walck 1984), which represents the upper mantle under a spreading centre, as only a minor part of the wave path is spent under such a structure. Model TNA is similar to GCA and is listed in table II with a reasonable energy content of the  $P$  coda, but this is due to the relatively low amplitude  $PP$ -waves. For this event all models fail to generate a  $PP$ -coda level that is even near the large amplitudes seen in the bottom trace in figure 8.

The results for the  $PP$  coda are not very stable. The three events are all within the  $PP$ -triplication distance, causing a strong dependence on upper mantle structure of the  $PP$ -wave train. The  $PP$  coda is a  $PP$ -coupled  $PL$  mode (see, e.g. Alsop & Chander (1968)); the reverberative characteristics of  $PP$ - $PL$  can be seen most clearly for event 2, see figure 3. It depends strongly on the crust and upper mantle above the LVZ near the receiver, as it is a leaky mode of this waveguide. This region of the upper mantle exhibits the strongest lateral variations in structure, which is the reason that the  $PP$ -waves and  $PP$  coda do not match observed  $PP$  and  $PP$  coda wave trains. Table II shows that PEMc is always among the models that predict a  $PP$  coda that is closest to the data with respect to its energy content; therefore PEMc also best represents the average structure of the upper mantle above the LVZ below GRF.

Both the  $P$  and  $PP$  coda observed in the data suggest an upper mantle with a thick lid over a more or less pronounced LVZ between GRF and Hindu Kush. This is in agreement with the results of Rial et al. (1984), who modeled  $S$ - and  $SS$ -waves for almost the same path as the present great circle path. They found that SNA modeled the data well for predominantly shield paths, whereas TNA gave satisfactory results for tectonic paths. The boundary between these regions appeared to be sharp. Paulssen (1988a) constructed a  $P$ - and  $S$ -velocity model of western Europe for events in the Mediterranean recorded by stations in Europe. and obtained a model also similar to SNA: it contains a thick lid over a pronounced LVZ. Given & Helmberger (1980), using Soviet nuclear explosions recorded in western Europe, constructed model K8 from wave form modeling. K8 has a relatively thin LVZ; although it generates a reasonable  $P$  coda, it hardly generates  $PP$  coda at all. The difference of the models K8 and PEMc is mainly the thickness of the LVZ.

It can thus be concluded that as far as  $P$  coda is concerned, no scattering is necessary to

explain the data, at these frequencies and at this distance. The synthetics contain a coda comparable to the data; detailed characteristics are not explained. The one exception is event 4, that shows a *P* coda level larger than explained by any of the models used here, but this is also due to the *PcP* phase not included in the synthetics. With the synthetics presented here, scattering at or near the source cannot be discarded, as waves scattered near the source appear as coherent waves near the receiver at suitably large distances. Greenfield (1971) showed that, for short-period *P* coda, Rayleigh waves scattered and converted to body waves by topography near the source could account for the observed decay of the coherent part of the coda. This may also account for the generally somewhat higher *P* coda level in the data.

## 5. Upper mantle discontinuities.

Although it was not our intention to model individual phases, some constraints on upper mantle discontinuities can be obtained from a more detailed comparison of the synthetics and data in figures 6, 7 and 8. Figure 6 shows that several models predict a pronounced  $P_{400}P$  phase at this epicentral distance. In the case of models 1066B and PEMc it is responsible for a considerable part of the energy of the *P* coda. Such large amplitudes are not found in the data, indicating that in these models the discontinuities are probably too large. Even model K8, with a moderate velocity jump at 400 km and derived for a region near the present great circle path, generates a rather large  $P_{400}P$  phase, which is reproduced by WKBJ synthetics. Possible reasons for the small amplitude in the data include: the discontinuity can be weaker than in the model K8, which explains why bottom-side reflections are rarely reported in the literature (see e.g., Wajeman 1988; Davis et al. 1989; Shearer 1990); the discontinuity exhibits lateral changes in strength (evidence for a laterally changing strength of the 670 km discontinuity is presented by Paulssen (1988b)); the interface may be curved, causing focusing and defocusing of the  $P_dP$  phase (Davis et al. 1989, see also chapter 4); other arrivals such as reverberations may mask the  $P_dP$  phase; the synthetic phase may be in a caustic while the data are not; the discontinuity is not a first-order contrast, but rather a gradient with a thickness comparable to the wavelengths considered here. WKBJ synthetics show that for model K8 the precursor is not in a caustic, although it is within the triplication zone of the 670-km discontinuity. Ingate et al. (1986) show with numerical experiments that a velocity gradient with wavelength longer than about a quarter of the seismic wavelength is no longer 'seen' as an abrupt velocity jump. The low amplitude of the observed (rather: the unobserved)  $P_{400}P$  phase may then also be explained by the 400-km discontinuity being a gradient rather than a sharp contrast. This gradient would then have to have an extent of about 20-25 km, for a dominant frequency of 0.1 Hz.

Some models contain a strong 200 km discontinuity, like PREM and PEMc. The synthetic seismograms for these models show a strong pulse between the  $P_{400}P$  and *PP* phases. Such a pulse is not observed in the data, and it is likely that these models overestimate the strength of the 200 km discontinuity, which would be in agreement with the result of Nolet (1990) for the *S*-velocity contrast at this depth in western Europe. At these epicentral distances, no constraints can be put on the 670 discontinuity, as the  $P_{670}P$  phase arrives within the *P* - *pP* - *sP* wave train.

In figure 6 the data of station B<sub>3</sub> show a strong phase arriving at about  $t=480$  s, which is coherent across the array (figure 3). It arrives too late to be a  $P_{400}P$  phase. Assuming that

it is a  $P_dP$  phase, it corresponds with  $d$  about 300 km. This depth was also found by Wajeman (1988), for  $P_dP$  phases from events near Japan recorded by the NARS array. The bounce point for this event is indicated with the open square in figure 1; it is near the Black Sea in the Caucasus. The results of Dercourt et al. (1986) indicate that subduction has been taking place for the last 200 Ma near this region. Remnants of subducted material present in the upper mantle may cause anomalous phases as this.

## 6. Sensitivity to LVZ structure.

The conclusion that models with a LVZ (in both  $P$ - and  $S$ -velocity) show a better fit of the data than models without a LVZ suggests the possibility of using the  $P$  coda to determine the presence or absence of a LVZ. The main arrivals associated with the LVZ are top- and bottom-side reflections at the top and base of the LVZ. Internal reverberations (peg-leg multiples) in the LVZ have small amplitudes, about 0.1% of that of  $P$ .

To determine the sensitivity of the  $P$  coda to LVZ-structure, synthetic seismograms are computed for a number of models. PEMc is used as starting model, it has a LVZ in both  $P$ - and  $S$ -velocity. Models with no LVZ or a LVZ in only one of  $P$ - and  $S$ -velocity are constructed from it. The synthetics are computed with the reflectivity method, using an isotropic source.

The synthetic seismograms are plotted in figure 9, using the synthetic for the original PEMc as reference. The traces have been lined up at the  $P$ -wave for comparison. No instrument response is applied; the traces are low-passed with corner frequency 0.15 Hz. Trace a in figure 9 shows the interval in the  $P$  coda where information from the LVZ arrives. The solid line is the total response, the dashed line represents the wave field when only transmission through the LVZ is allowed. Top- and bottomside reflections on the 220-km discontinuity, which marks the bottom of the LVZ, show up clearly. The interval with low amplitudes is the interval containing information about the LVZ. Trace b in figure 9 shows the result for model PEMc without the LVZ in the  $S$ -velocity (solid line; dashed line is the total response of the original PEMc). There are minor differences between the two traces, indicating low sensitivity of the  $P$  coda for the  $S$ -velocity structure. Removing the LVZ in the  $P$ -velocity has a significant effect on the synthetics (trace c in figure 9, solid line is PEMc without a LVZ in  $P$ -velocity, dashed line is original PEMc). Trace d of figure 9, finally, the synthetic for PEMc with velocities in the LVZ lowered by 1 km/s is shown (solid line, dashed is original PEMc). Due to what is probably interference between the  $P_{400}P$  and  $Pp200p$  (a reverberation on the 200-km discontinuity) phases, there are large amplitudes around 520 s.

The results in figure 9 show that the long-period  $P$  coda is sensitive to the  $P$ -velocity structure in the LVZ. Changes in the  $S$ -velocity hardly affect the synthetics, indicating that  $P-S$  converted waves in the coda have small amplitude. This suggests the possibility of using the phases in the  $P$  coda to monitor the velocity structure of the LVZ. Helmberger (1972) and York & Helmberger (1973) used the  $PL$  phase in the coda of  $P$  to study lateral variations of the LVZ in the United States. The  $P_dP$  phases used in this study have the advantage that one is not restricted to events within about  $15^\circ$  to  $20^\circ$  from the stations and that one can investigate regions on Earth where no earthquakes occur or no stations are located. In chapter 4  $P_{400}P$  phases are used to probe the upper mantle below the Hawaii hotspot region.

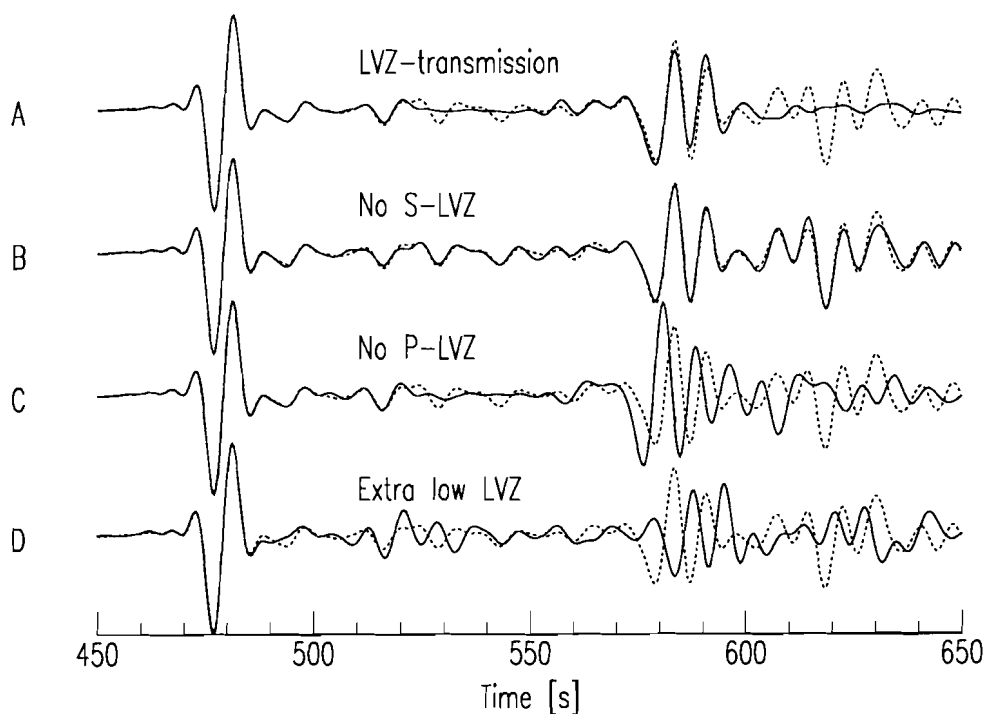


Figure 9. Reflectivity synthetics for model PEMc to test the sensitivity of the coda of *P* for LVZ-structure. In all traces the dashed line represents the total response of the original model PEMc. A) Solid: PEMc with only transmission of energy through LVZ. B) Solid: PEMc, no LVZ in *s*-velocity. C) Solid: PEMc, no LVZ in *P*-velocity. D) Solid: PEMc, velocities in LVZ lowered by 1 km/s.

## 7. Discussion.

The synthetic seismograms presented here prove that laterally homogeneous models of the upper mantle can explain the long-period *P* coda at teleseismic distances. The scattering mechanism that is used to explain short-period *P* coda characteristics (such as amplitude decay, incoherency) is not needed to explain the long-period coda; it is also inconsistent with the observed coherency of the the long-period *P* coda at GRF. The NARS array, on the other hand, shows hardly any coherency between two adjacent stations, even for stations at almost the same epicentral distance (see figure 4). Therefore, it can be concluded that the great circle paths from the events to the different NARS stations sample the lateral heterogeneity of the upper mantle under Europe. The GRF array is apparently too small. The scale length of these lateral variations must be comparable to or larger than the average station separation of NARS. This constrains the scale lengths of the lateral heterogeneity of the upper mantle to be one hundred to several hundreds of kilometers.

Combining the results presented here with earlier studies of the upper mantle under Europe suggests that the LVZ may be the origin of these lateral variations. The great circle paths studied by Given & Helmberger (1980) and Rial et al. (1984) are closely spaced; the

models proposed differ mainly in the thickness of the LVZ. Rial et al. further show that the boundary between tectonic and shield areas in Europe is probably quite sharp. These two areas are well modeled by the models TNA and SNA, respectively. The most prominent difference between these models is the absence in TNA of the well-pronounced LVZ of SNA. The results of Nolet (1990) show that the LVZ in  $S$ -velocity that is present under the Iberian peninsula disappears beneath the Paris Basin and the Baltic shield. Snieder (1988) used a surface wave scattering formalism and found significant ( $\pm 5\%$ ) lateral variations at a depth of 100-200 km under Europe.

Although there are some upper mantle models that come close to the data as far as the energy content of the  $P$  coda is concerned, there is one characteristic that is matched by none of the synthetics. Figure 7 most clearly shows the reverberative character of the long-period  $P$  coda. Whereas the synthetics have intervals of relatively low amplitude in between impulsive arrivals, there are no such pulse-like arrivals in the data. Again the explanation is the lateral heterogeneity of the upper mantle. The synthetics have been computed for identical structure beneath source and receiver. In this case reverberations have a multiplicity of more than one, and all equivalent reverberations (with the reverberation under either source or station) arrive at the same time. If the source area structure differs from the receiver area structure, the multiplicity is lost and the reverberations can set up a reverberative  $P$  coda. This will also make it difficult to pick impulsive arrivals, such as  $P_dP$  phases, as they are masked by the reverberations. When studying the  $P$  coda, one actually deals with three upper mantle structures, which are generally not identical: the structures beneath source and receiver and the upper mantle at the bounce point of the  $P_dP$  phases. Resolving these three upper mantle structures is difficult, if not impossible with only a few source-station combinations. For the present problem the assumption of identical structures is definitely not correct. However, some constraints on what represents an average upper mantle structure for the great circle paths studied can be found. That this is not an unreasonable assumption is shown by the agreement for the three events studied and by the agreement with this average structure and results reported in the literature for comparable great circle paths.

For the great circle paths studied here, the constraints that are found from synthetic seismogram analysis are a thick lid over a more or less pronounced, thick LVZ; models that do not have these features cannot explain the observed character (energy content) of the data. This suggests that the long-period  $P$  coda can be used to prove the existence or absence of a LVZ. Synthetic seismograms show that the  $P$  coda is more sensitive to  $P$ - than to  $S$ -velocity of the LVZ. The amplitude differences in the  $P$  coda correspond to topside and bottomside reflections at the bottom of the LVZ. Measuring the energy in the time interval in which arrivals from the LVZ are expected should make it possible to study the LVZ. Whereas the existence of a LVZ in the  $S$ -velocity layer can be proven from surface wave dispersion, the LVZ in  $P$ -velocity has not been proven unequivocally. Proving its existence or absence has important bearings on the structure and composition of the upper mantle. Estimates of the sharpness of the lower boundary of the LVZ can discriminate between theories concerning its origin, either partial melt, decreasing both  $S$ - and  $P$ -velocity (e.g. Anderson & Sammis 1970; Lambert & Wylie 1970; Stocker & Gordon 1975) or a simple pressure effect, causing the  $S$ -velocity to decrease but the  $P$ -velocity to increase (Tolstoy 1982).

One of the possible reasons for the null observation of the  $P_{400}P$  phase in the data is that

the 400-km discontinuity is not an abrupt velocity increase, but a gradient of about 20-25 km thick, for a dominant period of 10 s. Although this is only one of several explanation for the null observation of this phase, it is interesting to compare this value with previously reported values for the thickness of the discontinuity. It compares well with high-pressure and high-temperature laboratory experiments from Katsura & Ito (1989), who estimate the velocity to increase over an depth interval of about 10-20 km, on the assumption that the discontinuity is caused by the olivine-modified spinel phase transition. From slowness and travel time analysis of short-period body wave data, Fukao (1977) estimated it to be 5-15 km, whereas Leven (1985) claimed a thickness of about 5 km under northern Australia.

## References.

- Aki, K., 1973. Scattering of *P*-waves under the Montana Lasa, *J. Geophys. Res.*, **78**, 1335-1346.
- Aki, K., 1980a. Attenuation of shear waves in the lithosphere for frequencies from 0.05 to 25 Hz, *Phys. Earth Planet. Int.*, **21**, 50-60.
- Aki, K., 1980b. Scattering and attenuation of shear waves in the lithosphere, *J. Geophys. Res.*, **85**, 6496-6504.
- Alsop, L.E. & R. Chander, 1968. The generation of direct and coupled *PL* modes, *Suppl. Nuovo Cimento*, **6**, 7-21.
- Anderson, D.L. & Sammis, C., 1970. Partial melting in the upper mantle, *Phys. Earth Planet. Int.*, **3**, 41-50.
- Baag, C.-E. & C.A. Langston, 1985a. Shear-coupled *PL*, *Geophys. J. R. astron. Soc.*, **80**, 363-385.
- Barley, B.J., J.A. Hudson & A. Douglas, 1982. *S* to *P* scattering at the 650-km discontinuity, *Geophys. J. R. astron. Soc.*, **69**, 159-172.
- Bolt, B.A., 1970. *P<sub>d</sub>P* and *PKiKP* waves and diffracted *PcP* waves, *Geophys. J. R. astron. Soc.*, **20**, 367-382.
- Burdick, L.J., 1981. A comparison of the upper mantle structure beneath North America and Europe, *J. Geophys. Res.*, **86**, 5926-5936.
- Chapman, C.H. & J.A. Orcutt, 1985. The computation of body wave synthetic seismograms in laterally homogeneous media, *Rev. Geophys.*, **23**, 105-163.
- Davis, J.P., R. Kind & I.S. Sacks, 1989. Precursors to *P'P'* re-examined using broad-band data, *Geophys. J. Int.*, **99**, 595-604.
- Der, Z.A., A.C. Lees & V.F. Cormier, 1986. Frequency dependence of *Q* in the mantle underlying the shield areas of Eurasia, part III: the *Q* model, *Geophys. J. R. astron. Soc.*, **87**, 1103-1112.
- Dercourt, J., L.P. Zonenshain, et al., 1986. Geological evolution of the Tethys belt from the Atlantic to the Pamirs since the Lias, *Tectonophysics*, **123**, 241-315.
- Dziewonski, A.M. & D.L. Anderson, 1981. Preliminary reference Earth model, *Phys. Earth Planet. Int.*, **25**, 297-356.
- Dziewonski, A.M., A. Hales & E Lapwood, 1975. Parametrically simple Earth models consistent with geophysical data, *Phys. Earth Planet. Int.*, **10**, 12-48.
- Flatté, S.M. & R.-S. Wu, 1988. Small-scale structure in the lithosphere and asthenosphere deduced from arrival time and amplitude fluctuations at NORSAR, *J. Geophys. Res.*, **93**, 6601-6614.
- Frankel, A. & L. Wennerberg, 1987. Energy-flux model of seismic coda: separation of scattering and intrinsic attenuation, *Bull. Seismol. Soc. Am.*, **77**, 1223-1251.
- Fukao, Y., 1977. Upper mantle *P* structure on the ocean side of the Japan-Kurile arc, *Geophys. J. R. astron. Soc.*, **50**, 621-642.
- Gilbert, F. & A.M. Dziewonski, 1975. An application of normal mode theory to the retrieval of structural parameters and source mechanisms from seismic spectra, *Phil. Trans. R. Soc. London*, **A278**,

- 187–269.
- Given, J.W. & D.V. Helmberger, 1980. Upper mantle structure of northwestern Eurasia, *J. Geophys. Res.*, **85**, 7183–7194.
- Grand, S.P. & D.V. Helmberger, 1984. Upper mantle shear structure of North America, *Geophys. J. R. astron. Soc.*, **76**, 399–438.
- Greenfield, R.J., 1971. Short-period *P*-wave generation by Rayleigh-wave scattering at Novaya Zemlaya, *J. Geophys. Res.*, **76**, 7988–8002.
- Gudmundsson, O., Davies, J.H. & Clayton, R.W., 1990. Stochastic analysis of global traveltimes: mantle heterogeneity and random errors in the ISC data, *Geophys. J. Int.*, **102**, 25–43.
- Gutowski, P.R. & E.R. Kanasevich, 1974. Velocity spectral evidence of upper mantle discontinuities, *Geophys. J. R. astron. Soc.*, **36**, 21–32.
- Helmberger, D.V., 1972. Long-period body wave propagation from 4° to 13°, *Bull. Seism. Soc. Am.*, **62**, 325–341.
- Helmberger, D.V. & G.R. Engen, 1980. Modeling the long-period body waves from shallow earthquakes at regional ranges, *Bull. Seismol. Soc. Am.*, **70**, 1699–1714.
- Husebye, E.S. & R. Madariaga, 1970. The origin of precursors to core waves, *Bull. Seismol. Soc. Am.*, **60**, 939–952.
- Husebye, E., R.A.W. Haddon & D.W. King, 1977. Precursors to *P'P'* and upper mantle discontinuities, *J. Geophys.*, **43**, 535–543.
- Ingate, S.F., J. Ha & K.J. Muirhead, 1986. Limitations on waveform modelling of long-period seismograms, *Geophys. J. R. astron. Soc.*, **86**, 57–61.
- Katsura, T. & E. Ito, 1989. The system  $Mg_2SiO_4 - Fe_2SiO_4$  at high pressures and temperatures: precise determination of stabilities of olivine, modified spinel and spinel, *J. Geophys. Res.*, **94**, 15663–15670.
- Kennett, B.L.N., 1987. Observational and theoretical constraints on crustal and upper mantle heterogeneity, *Phys. Earth Planet. Int.*, **47**, 319–332.
- Kennett, B.L.N., 1988. Systematic approximations to the seismic wave field, in *Seismological Algorithms*, edited by D.J. Doornbos, pp 237–259, Academic, London.
- Kind, R. & L.P. Vinnik, 1988. The upper mantle discontinuities underneath the GRF array from *P*-to-*S* converted phases, *J. Geophys.*, **62**, 138–147.
- King, D.W., R.A.W. Haddon & E.S. Husebye, 1975. Precursors to *PP*, *Phys. Earth Planet. Int.*, **10**, 103–127.
- Korn, M., 1988. *P*-wave coda analysis of short-period array data and the scattering and absorptive properties of the lithosphere, *Geophys. J.*, **93**, 437–449.
- Lambert, I.B. & Wylie, P.J., 1970. Low-velocity zone of the Earth's mantle: incipient melting caused by water, *Science*, **169**, 764–766.
- Leven, J.H., 1985. The application of synthetic seismograms to the interpretation of the upper mantle *P*-wave velocity structure in northern Australia, *Phys. Earth Planet. Int.*, **38**, 9–27.
- Mallick, S. & L.N. Frazer, 1987. Practical aspects of reflectivity modeling, *Geophysics*, **52**, 1355–1364.
- Nakanishi, I., 1988. Reflections of *P'P'* from upper mantle discontinuities beneath the Mid-Atlantic ridge, *Geophys. J.*, **93**, 335–346.
- Nolet, G., 1990. Partitioned wave form inversion and 2D structure under the NARS array, *J. Geophys. Res.*, **95**, 8499–8512.
- Oliver, J. & M. Major, 1960. Leaking modes and the *PL* phase, *Bull. Seimol. Soc. Am.*, **50**, 165–180.
- Paulssen, H., 1985. Upper mantle converted waves beneath the NARS array, *Geophys. Res. Lett.*, **12**, 709–712.
- Paulssen, H., 1988a. Lateral heterogeneity of Europe's upper mantle as inferred from modelling of broad-band body waves, *Geophys. J. R. astron. Soc.*, **91**, 171–199.
- Paulssen, H., 1988b. Evidence for a sharp 670-km discontinuity as inferred from *P* to *S* converted waves, *J. Geophys. Res.*, **93**, 10489–10500.



- Rial, J.A., S. Grand & D.V. Helmberger, 1984. A note on lateral variation in upper-mantle shear velocity across the Alpine front, *Geophys. J. R. astron. Soc.*, **77**, 639–654.
- Shearer, P., 1990. Seismic imaging of upper mantle structure - new evidence for a 520 km discontinuity, *Nature*, **344**, 121–126.
- Snieder, R., 1988. Large scale waveform inversions of surface waves for lateral heterogeneity - II: Application to surface waves in Europe and the Mediterranean, *J. Geophys. Res.*, **93**, 12067–12080.
- Spakman, W., 1991. Delay time tomography of the upper mantle below Europe, the Mediterranean, and Asia Minor, *Geophys. J. Int.*, **107**, 309–332.
- Stocker, R.L. & Gordon, R.B., 1975. Velocity and internal friction in partial melts, *J. Geophys. Res.*, **80**, 4828–4836.
- Tolstoy, I., 1982. On elastic waves in prestressed solids, *J. Geophys. Res.*, **87**, 6823–6827.
- Toksöz, M.N., A.M. Dainty & E.E. Charette, 1989. Coherency of ground motion at regional distances and scattering, *Phys. Earth Planet. Int.*, submitted.
- Wajeman, N., 1988. Detection of underside *P* reflections at mantle discontinuities by stacking broadband data, *Geophys. Res. Lett.*, **15**, 669–672.
- Walck, M.C., 1984. The *P*-wave upper mantle structure beneath an active spreading centre: the gulf of California, *Geophys. J. R. astron. Soc.*, **76**, 697–723.
- Whitcomb, J.H. & D.L. Anderson, 1970. Reflection of *P*'*P*' seismic waves from discontinuities in the mantle, *J. Geophys. Res.*, **75**, 5713–5728.
- York, J.E. & D.V. Helmberger, 1973. Low-velocity zone variations in the southwestern United States, *J. Geophys. Res.*, **78**, 1883–1886.

## Chapter 3

### Are long-period body-wave coda caused by lateral heterogeneity? - II. The coda of PP

**Summary.** The long-period coda of *PP*-, *PPP*- and *S*-waves is shown to be incoherent at the Gräfenberg array (GRF), which has an average station separation of about 10 km. This indicates that these coda consist of scattered waves. The data of GRF are analyzed for slowness and azimuth variations in the coda intervals. A new beamforming algorithm is presented to deal with the low frequencies and relatively short time intervals, associated with long-period body-wave data. The method is based on Backus-Gilbert inverse theory. The results show that the long-period coda intervals almost entirely consist of surface waves; these waves are scattered from the preceding body waves. Some calculations with linearized theory for body wave to Rayleigh wave conversion at topography at the surface or at the Moho show that realistic scatterers can account for the observed (constant) coda level.

#### 1. Introduction.

This chapter deals with the long-period coda of *PP/PPP*- and *S*-waves. The aim of this and the previous chapter is to determine the class of models that are consistent with long-period body wave data, and to find out whether different parts of the seismogram require different classes of Earth models.

The results in the chapter 2 indicate that the teleseismic *P* coda can be explained by spherically symmetric Earth models. No scattering is needed to explain the character of this coherent part of the seismogram. In contrast, the coda following *PP*- and *S*-waves is incoherent at a small scale, and cannot be explained by wave propagation in a laterally homogeneous Earth model. The incoherent behaviour suggests that scattering is important in these parts of the seismogram. In this chapter it is studied which wave scattering mechanisms are responsible for the incoherent behaviour of the coda intervals following *PP*- and *S*-waves.

The scattering approach has been very successful in explaining short-period coda waves. Coda waves have been interpreted as waves scattered either singly or multiply in a medium

with random velocity fluctuations. Many studies have been devoted to the determination of statistical parameters describing the random medium (e.g., Aki 1973; Kennett 1987; Flatté & Wu 1988; Korn 1988). Interpreting coda waves from local events as backscattered waves may yield estimates of the scale lengths present (Aki & Chouet 1975; Gao et al. 1983; Sato 1984). Lateral heterogeneity in an otherwise plane layered or homogeneous medium also produces substantial coda. Hill & Levander (1984) and Levander & Hill (1985) showed that random perturbations in the depth to the interface between a layer and a half-space causes strong coupling between incident body waves and the guided modes of the layer. Lateral variations in layer thickness also produces this coupling (Fuyuki & Matsumoto 1980; Frankel & Clayton 1986; Stead & Helmberger 1988). Greenfield (1971) explained observed short-period  $P$  coda decay by assuming that the coda waves were due to Rayleigh-to- $P$  coupling at topography near the source. Another coda generating mechanism should be mentioned here, namely that of waves reverberating in a plane layered structure. Bouchon (1982) modeled complete seismograms from local events by summing all crustal reverberations. Sereno & Orcutt (1988) explained the coda of  $P_n$  and  $S_n$  as reverberations. The long-period  $PL$  mode is a reverberation of the crust (Helmberger & Engen 1980; Baag & Langston 1985). When the reflection is specular, a long coda may be built up. In a laterally homogeneous medium such coda waves will be coherent, but if there are variations in the depth to interfaces this coherence will be lost. Palmer et al (1988) showed that small lateral variations in, for example, Moho depth results in chaotic behaviour of ray paths of reverberating waves, resulting in completely incoherent coda waves. Bostock & Kennett (1990) computed  $Lg$  ray paths through a laterally varying crustal structure and indeed found that wave paths behave erratically. Toksöz et al. (1990) studied the influence of random interface depth perturbations (and random velocity perturbations) on the coherency of short-period  $Lg$  phase and matched observed features of the coherency.

Summarizing, it seems that scattering is an efficient mechanism to generate coda waves, especially at short periods. The scattering may occur at random velocity perturbations or at lateral variations of the otherwise plane layered structure, such as depth to the interfaces or tilted interfaces. Coupling between different wave types is expected to occur, especially in the latter case. Repeated specular reflection in a waveguide can also build up considerable coda, and if the waveguide varies laterally this coda will be incoherent.

To find constraints on possible scattering mechanisms a new beamforming algorithm is developed and used to study slowness and azimuth variations throughout the body wave part of the seismogram. Slowness and azimuth variations should give some clues as to the nature of the scattering. It is shown that the coda following  $PP/PPP$ - and the  $S$ -waves consists of surface waves, scattered from the preceding body waves at lateral heterogeneity in the lithosphere. (In this context, the term scattering also includes mode conversion; in the process of conversion also the direction of propagation can be different from that of the incident wave.) Finally some calculations using linearized theory are performed to show that realistic scatterers (such as topography at the surface or at the Moho) can explain the observed coda characteristics, such as duration, amplitude and slowness content. The data are from the broad-band GRF and NARS arrays and from events at  $40^\circ$  to  $60^\circ$  distance.

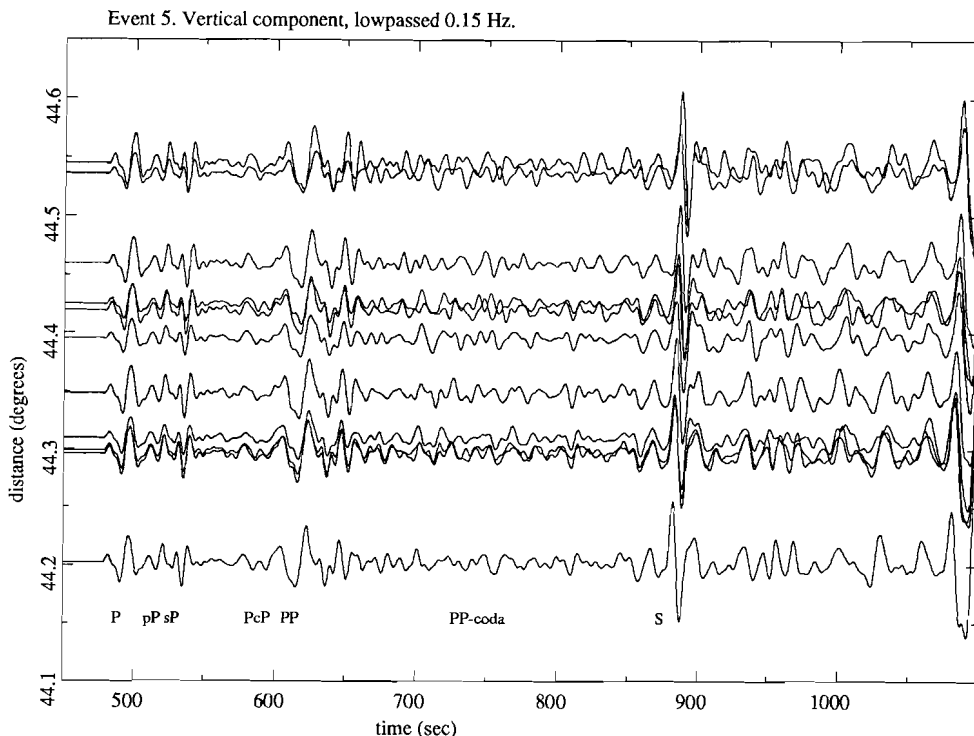


Figure 1. GRF data from event 5. Vertical components, lowpassed at 0.15 Hz. Although the coda of  $P$  is dominated by  $pP$ ,  $sP$  and  $PcP$ , the coherency can be seen. The coda of  $PP$  is largely incoherent. The coda of  $S$  also contains many incoherent arrivals.

## 2. Data.

As in the previous chapter, the data have been taken from two broad-band arrays, GRF and NARS. The events used here are listed in table I of chapter 2. The reader is referred to figures 1 and 2 of chapter 2 for the configuration of the arrays and events.

Figure 1 shows the data for event 5; the data are from the GRF array. The data have been lowpassed with corner frequency 0.15 Hz. Note the large difference in coherency between the  $P$  coda and the coda of  $PP$ - and the  $S$ -waves. The interval between  $pP$  and  $sP$  and  $PcP$  is strongly coherent. This is in agreement with the results of chapter 2. In contrast, the coda of  $PP/PPP$  is incoherent, indicating that in this interval scattering is important. This sharp contrast is found for all events studied here: a strongly coherent  $P$  coda, in which nearly all waveforms can be traced from one station to the next, and coda intervals following  $PP$ - and  $S$ -waves that are dominated by incoherent energy. Chapter 2 contains more examples, that also show the different character of NARS data. The NARS data show incoherency of the  $P$  coda, as well as of the other coda intervals. It is shown in Chapter 2 that this difference in coherency of the  $P$  coda at the two arrays can be explained by large-scale lateral variations in an otherwise spherically symmetric upper mantle. The scale length of these variations must be several hundred kilometers. The coda of  $PP$ - and  $S$ -waves are incoherent at both

arrays. In the following only the data from GRF are analyzed to find the mechanism(s) of wave scattering.

### 3. Slowness and azimuth analysis.

Primary indicators for lateral heterogeneity are the slowness and azimuth of the seismic signal. In a transversely isotropic Earth, any deviations from great circle azimuth must be caused by aspherical structure. An array of stations can be used to measure slowness and azimuth by stacking the individual traces, each delayed or advanced with the proper time shift. This can be done in both time and frequency domain and is known as beamforming. Using this technique, Bungum & Capon (1974) and Levshin & Berteussen (1979) studied surface wave coda recorded at the NORSAR array and found surface waves possibly scattered by continental margins. Capon (1970) used array data of the LASA in Montana and found similarly refracted Rayleigh waves.

A number of beamforming techniques exist in the literature. The simplest algorithm is to simply delay and sum the individual traces: this is conventional beamforming. Its performance can be improved if weights are assigned to each station. Capon (1969) developed a high resolution algorithm, where the weights are designed such as to optimize the array response. This technique was developed for application to surface wave problems and is not used in this study; this is due to the inherently short time series in body wave problems, which leads to instability of the high-resolution technique.

For the present study a new beamforming algorithm has been developed for applications to body wave problems. In order to optimize the array response at the long periods used here, a Backus-Gilbert inversion technique (BGI) (Backus & Gilbert 1968) was applied to the conventional (simple phased sum) beamforming. Appendix A gives an outline of the theory. This optimization does not suffer from the drawbacks of the high-resolution technique of Capon (1969), when applied to short intervals. The stacking is done in the frequency domain.

Only a limited frequency interval can be studied, due to the size of GRF. The frequency dependent Nyquist  $p_N$  slowness for an array with station separation  $D$  is  $p_N = (2Df)^{-1}$ . The slowness interval can be taken such that all physically possible slownesses are included. With the station separation  $D$  of 10 km of GRF and 0.35 s/km (2.86 km/s) as an upper limit on expected slownesses (see Dost (1990) for surface wave velocities in western Europe), this gives a frequency of 0.15 Hz below which the Nyquist criterion is satisfied. Frequencies higher than this are spatially aliased and artifacts will contaminate the wavenumber spectrum. A lower bound on the frequency interval is obtained by considering that waves with wavelengths greater than the extent of the array are not resolved very well. Thus, 0.03 Hz is a safe lower bound. The BGI is applied only to slownesses  $|p| < 0.35$  s/km.

The events in table I of chapter 2 were chosen to have a backazimuth near either 80° or 260°. In this way the best azimuthal resolution in the direction of the source is obtained, so that any deviations from great circle azimuth are optimally recognized. A moving window analysis is applied to the data, the length of the window (100 s) is chosen as a compromise between resolution in the time and frequency domain. The windows do not overlap. The frequency interval 0.03-0.15 Hz is split into four subintervals. For each subinterval, the traces are prefiltered and Hanning-tapered to avoid artifacts due to the finite window length

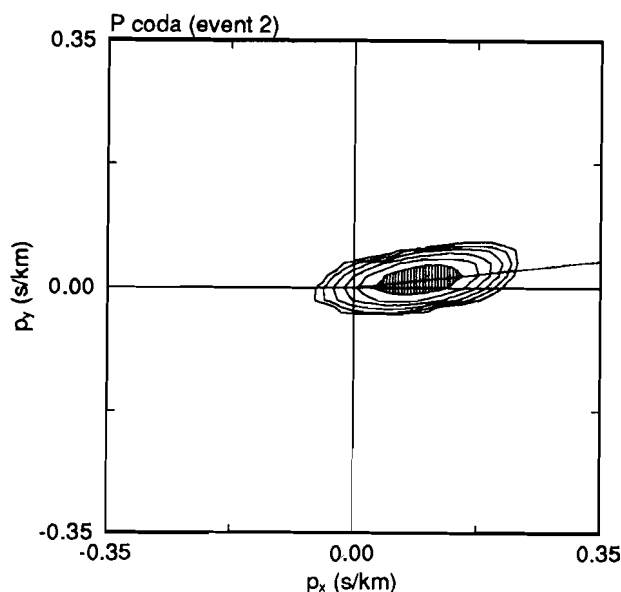


Figure 2. Example of a frequency-slowness spectrum. Line contours are in steps of 2 dB down from the maximum. The area above -2 dB is shaded; the spectrum is CLEANed. The spectrum is obtained from event 2,  $450 < t < 510$  s, frequency interval 0.10-0.13 Hz. This time interval contains the *P* coda. The back azimuth is indicated by the solid line in the direction of  $84^\circ$ . The *P* coda arrives on the great circle.

(see Capon 1970). Then the traces of all stations are Fourier-transformed. Erroneous station calibrations may influence the results; to prevent this the amplitude spectra are prewhitened. In this way, the information contained in the amplitudes is lost. To remove confusing side lobes from the slowness diagram the CLEAN algorithm (Högbohm 1974) is applied.

Figures 2,3 and 4 show three examples of slowness diagrams, obtained from event 2 and 5. The slowness diagram is plotted in such a way that a wave arriving from the east with slowness 0.10 s/km appears in the plot at  $p_x = 0.10$  and  $p_y = 0$  s/km. The direction of the great circle is indicated by the solid line. Contour lines are drawn at 2 dB intervals down from the maximum. Figure 2 shows the slowness diagram for the frequency band 0.12 - 0.15 Hz for a time interval containing the *P* coda of event 2. The *P* coda clearly arrives at great circle azimuth. Figure 3 is a slowness diagram for a time interval in the coda of *PP*, for the frequency band 0.09-0.12 Hz and the same event. Apart from some energy arriving along the great circle with body wave slowness, there are maxima off the great circle, with a large slowness. These are probably surface waves. As the signal-to-noise ratio for this event is very good, these surface waves must be signal-generated. Figure 4 is a slowness diagram obtained from the coda of the *S*-wave of event 5. The frequency interval is 0.06 - 0.08 Hz. There is no energy arriving along the great circle, but there are two surface waves arriving with azimuths deviations of about  $25^\circ$  and  $60^\circ$ .

The results of analyzing a large number of slowness diagrams are summarized in three figures. Figure 5 shows the absolute deviations from great circle azimuth as a function of time, of the most significant energy peaks in the slowness diagrams. The deviations are plotted at the position of the 100 s long, consecutive time intervals in the seismograms. These intervals are labeled with the body wave phase they contain (e.g. *P* or *S*), or with the

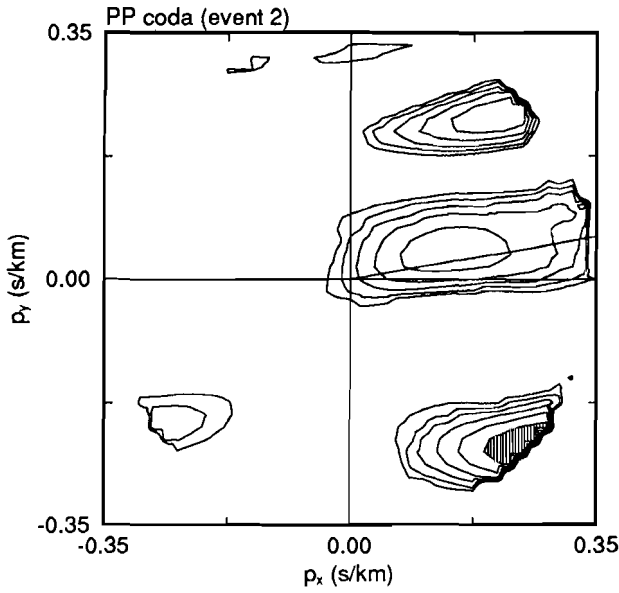


Figure 3. As figure 2, now for  $600 < t < 700$  s. This spectrum is in the coda of *PP*. Apart from some energy arriving on the great circle, there are two clear arrivals with slowness about  $0.34$  s/km and azimuth of  $40^\circ$  and  $131^\circ$ . The slowness of the waves indicates that these are surface waves.

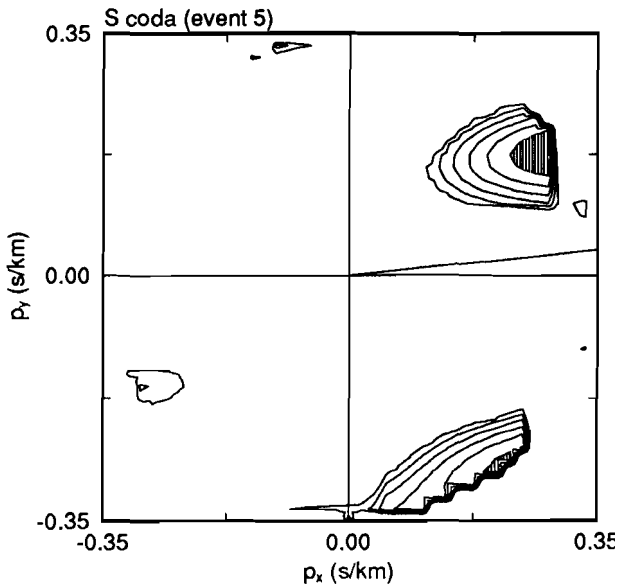


Figure 4. As figure 2, now for event 5, with  $950 < t < 1050$  s. This interval is in the coda of *S*. In this case there is no detectable energy arriving on the great circle; again surface waves arrive at azimuths of  $58^\circ$  and  $140^\circ$ .

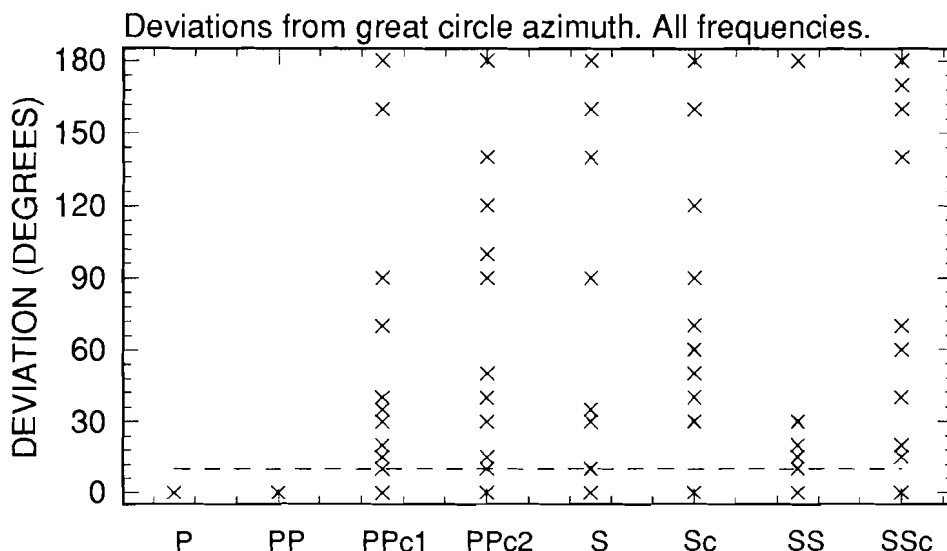


Figure 5. Absolute azimuth deviations of most prominent peaks in the slowness spectra. Data from all frequencies are plotted. The data are plotted as a function of the interval they were obtained from: *P*, *PP*, *S* etc. denote intervals containing these waves; labels with a suffix 'c' denote coda intervals. The dashed line indicates the resolution of GRF at a period of 10 s.

coda interval they represent (e.g. *PPc1* stands for the first interval in the *PP* coda). The dashed line represents an estimate of the resolution of GRF at a period of 10 s for events with back azimuth of either  $80^\circ$  or  $260^\circ$ . Waves arriving within  $10^\circ$  of the great circle are assumed to arrive on the great circle. Figure 5 contains data from all frequency bands. The data show a striking increase in deviation from the great circle after the *PP*-wave. Whereas *P* and *PP* do not show any deviation, the energy in the *PP* coda comes from all directions. A number of short time intervals in the coda of *P* have been analyzed as well, but no deviations were found. This is in agreement with the observation that the *P* coda is very coherent and the *PP* coda is not. The large deviations in the *PP* coda persist in the later intervals. Furthermore, the deviations do not seem to increase further into the coda.

The slowness of the coda wave can be used to identify the wave type. The slowness of only the energy peaks off the great circle is plotted in figure 6. The energy peaks corresponding to the direct waves are not considered in this plot. Again data from all frequency bands are plotted. The figure shows that apart from some waves with slownesses between 0.1 and 0.2 s/km, most of the scattered waves have a slowness larger than about 0.28 s/km. This corresponds with surface wave slownesses. Therefore, the scattered waves are mainly surface waves and they must have been scattered from body waves to arrive so early in the seismogram.

Figure 7 shows the frequency dependence of the direction of incoming of the scattered waves in the coda intervals. The great circle deviation data from all the coda intervals are combined and plotted as a function of frequency. It is clear from the figure that with increasing frequency the deviations from the great circle increase; this implies that



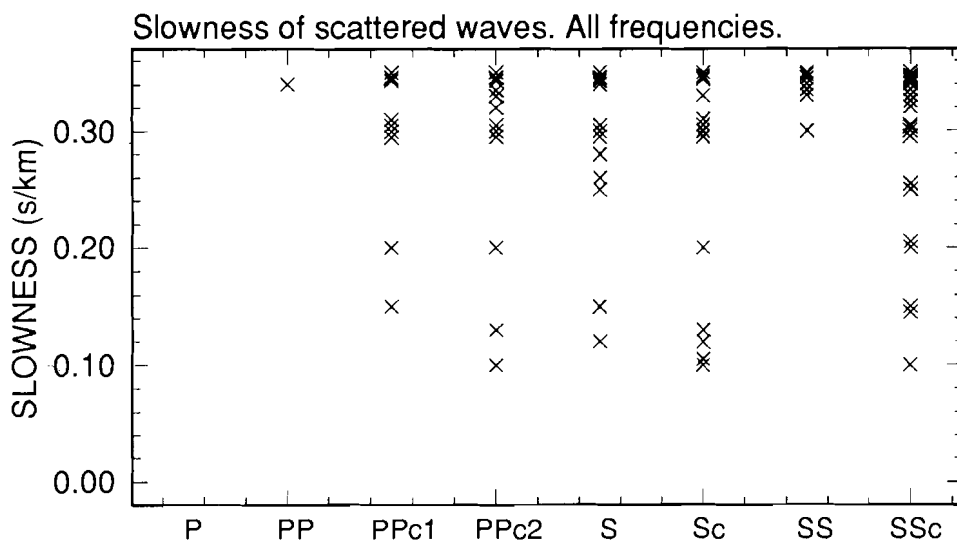


Figure 6. The slowness of the most prominent peaks in the slowness spectra, plotted in the same way as in figure 7. Note the concentration of data points at high slownesses, indicating that these waves are surface waves.

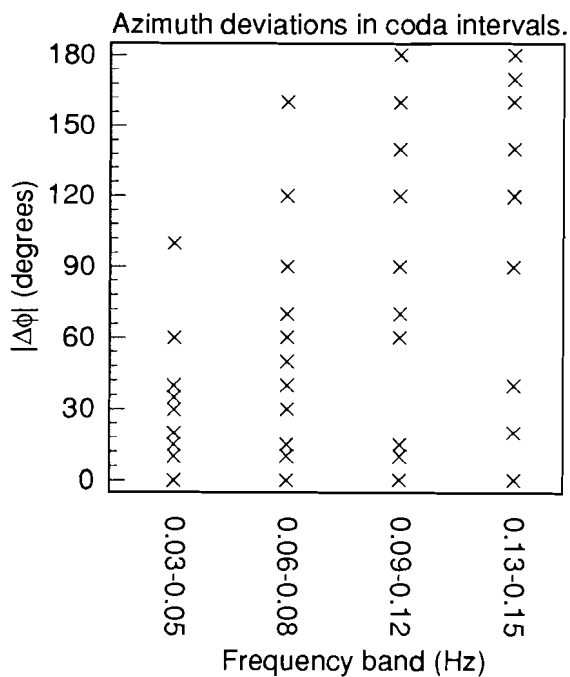


Figure 7. Absolute azimuth deviations as a function of frequency for data from coda intervals only. The deviations from back azimuth increase with increasing frequency.

backward scattering becomes more important at higher frequencies.

A number of conclusions can be drawn from the analysis of the slowness diagrams, even though the resolution of GRF is not very good. First, body waves do not show deviations from great circle azimuth, but their coda do. The one exception is the coda of *P*-waves, that shows much less scatter than the coda of *PP* and *S*-waves. This agrees well with the observation that this coda is very coherent across the array (see chapter 2). Secondly, scattering becomes more effective with increasing frequency. From scattering theory, this is to be expected, as scattering is a frequency-dependent process, the higher frequencies being scattered more efficiently. Thirdly, the amount of scatter does not increase further in the coda. This suggests a homogeneous coda in both space and time. This is also what is observed in the numerical calculations of Frankel & Wennerberg (1987), for the coda in a medium with random velocity fluctuations. These authors observed a coda level that was independent of epicentral distance and decayed only slowly. Finally, the slowness diagrams of coda intervals show much energy concentrated at surface wave slownesses. This indicates that surface waves make up an appreciable part of the coda.

In the next sections the frequency content of the coda waves is measured to obtain more information on the scattering process and possibly an estimate of the size of the scatterers.

#### 4. Frequency content of coda waves.

The frequency content of coda waves is a valuable source of information on scattering processes that produce the coda. Frequency losses due to absorption are ignored here. The frequency content of the *PP* coda is measured relative to that of the *PP*-waves by dividing the coda amplitude spectrum by the *PP*-wave amplitude spectrum. The latter is raised slightly to remove possible holes in the spectrum. This is then averaged over all stations of the array:

$$F(\omega) = \frac{1}{N} \sum_{i=1}^N \frac{PPc_i(\omega)}{PP_i(\omega) + \epsilon}, \quad (1)$$

where  $PP_i(\omega)$  and  $PPc_i(\omega)$  are the Fourier transformed *PP* and *PP* coda intervals of station *i*, respectively;  $\epsilon$  is a small number; *N* is the number of stations (which is 13 for GRF). In this way, the source and receiver effects are canceled.

Figure 8 shows two spectra of *PP* coda intervals, each divided by the *PP* amplitude spectrum. The spectra are the averages of the spectra of all 13 GRF stations. For most events, the spectra show low amplitudes for frequencies lower than about 0.2 Hz, whereas for higher frequencies the spectra are saturated. This is shown in figure 8<sup>a</sup>. One exception is event 4, which shows an exceptionally large *PP* coda level (figure 8<sup>b</sup>; for a time-distance plot see figure 5 of chapter 2).

The spectrum in figure 8<sup>a</sup> shows that the frequencies higher than about 0.2 Hz are scattered more efficiently than lower frequencies. It is reasonable to assume that the frequencies of interest in this study, about 0.03-0.15 Hz, are within the weak scattering regime. Event 4 is the exception to the rule. This deviating observation will be explained in a later section. For higher frequencies the spectra are saturated, indicating strong scattering. The frequency of 0.2 Hz is also the frequency at which microseisms have their peak energy. However, these will not influence the spectra, because the data have a very good S/N ratio (see figures 3 and 5 of chapter 2).

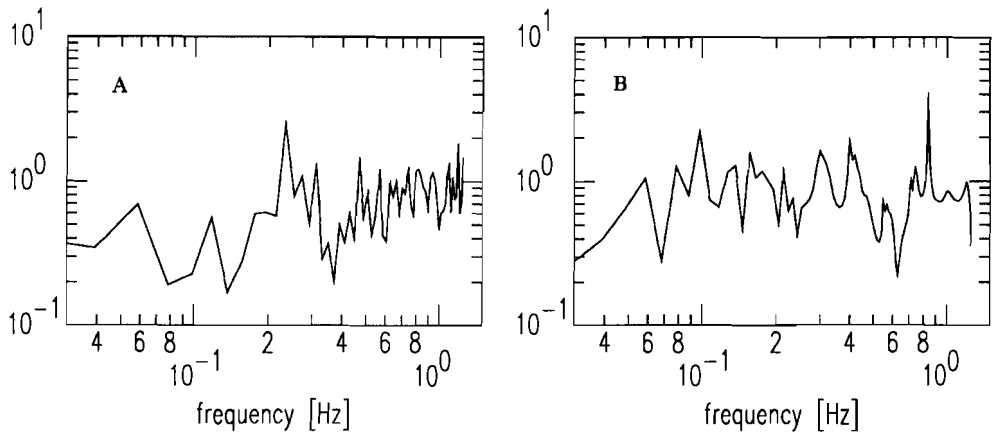


Figure 8. Spectra of *PP* coda intervals; the spectra are divided by the *PP* wave spectra. A: *PP* coda spectrum of event 2; B: *PP* coda spectrum of event 4.

An estimate can be given of the length scale of the scatterers, responsible for the increase in coda level at 0.2 Hz. Seismic waves with frequency of 0.2 Hz have a wave length of about 30-50 km in the crust and upper mantle. As scatterers with length scales greater than that of the seismic wave have small influence on the wave, the scatterers responsible for the saturation above 0.2 Hz must have length scales equal to or shorter than 30-50 km. Thus, the data indicate that the crust and upper mantle are rich in such scatterers. Note that this is much smaller than the cell size used in tomographic inversions (see, e.g., Spakman & Nolet 1988). The next sections deal with the problem of explaining the scattering at lower frequencies.

## 5. Body wave scattering.

Scattering is important in the *PP* and *S* coda, and it would be interesting to see what kind of scatterers are necessary to explain these coda amplitudes. Let us for the moment ignore the beamforming results and let our working hypothesis be that the coda waves are body waves scattered from body waves near source or receiver. This is in the spirit of King et al. (1975), who showed that the scattering of *P*-waves near source or receiver is an adequate explanation of observed slowness and azimuth variations in high frequency (above 0.6 Hz) *P* coda. The same mechanism may be at work at the lower frequencies studied here, and possibly in the *PP* and *S* coda as well as in the *P* coda. Assuming this mechanism, scatterer strength necessary to explain the observed coda level can be estimated. It is assumed that the Born approximation is valid: the waves are scattered only once. Elastic wave scattering theory of Wu & Aki (1985) is used to compute  $P \rightarrow P$  and  $P \rightarrow S$  scattered waves. The volume of the inhomogeneity is accounted for through the use of Mie scattering (which includes an integral over the inhomogeneity). The scattering structure is assumed to be a homogeneous sphere with constant perturbation in velocity or impedance. For a given lapse time, the scattered wave has been scattered somewhere on an ellipse, focused on source and receiver. The strength of a scattered wave varies along the ellipse, as the scattering amplitude depends on the scattering angle. It is assumed that the wave is scattered by

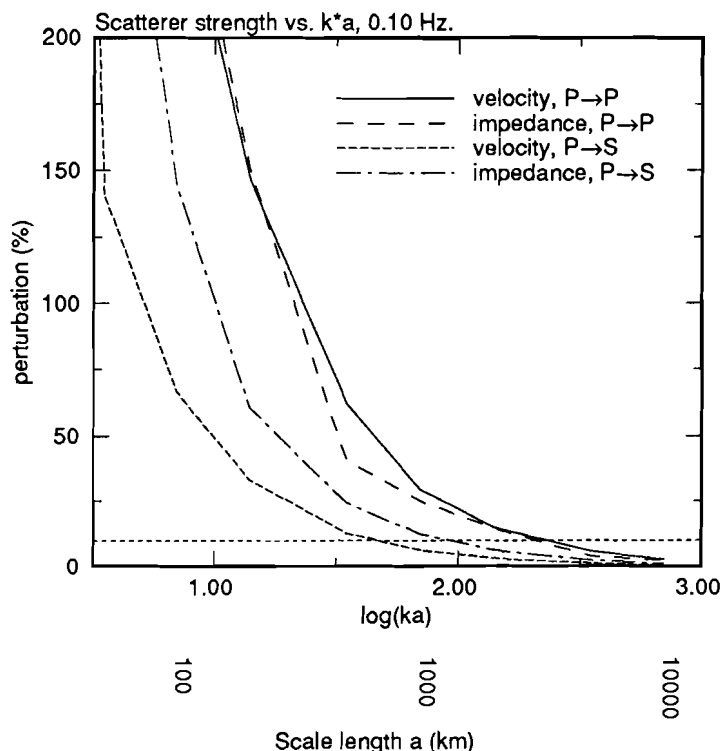


Figure 9. Strength of scatterer (either velocity or impedance perturbation) versus  $ka$ , seismic wavenumber times scatterer scale length, for a wave with frequency 0.1 Hz, lapse time 50 s and amplitude 20% of that of the direct wave. Epicentral distance is 4500 km. The scatterer is the most efficient scatterer on the ellipse focused on source and receiver, calculated using the theory of Wu & Aki (1985). Results are for  $P \rightarrow P$  and  $P \rightarrow S$  scattering. The dashed horizontal line represents a perturbation of 10%, which is assumed to be a reasonable value for inhomogeneities in the lithosphere.

the most efficient scatterer on this ellipse; in this way the most conservative estimate of scatterer strength is obtained.

Scatterer size and strength are computed to explain a wave with an amplitude of 0.2 times the amplitude of the preceding body wave and frequency 0.10 Hz (which is representative of the coda level in figure 1), and arriving with a 50 s time lag. The results are presented in figure 9, as either velocity perturbations (forward scattering) or impedance perturbations (backward scattering) as a function of the product  $ka$ , seismic wavenumber times scale length of the scatterer. The background velocity is taken to be 9 km/s. It turns out that either unacceptable perturbations ( $>100\%$ ) or unreasonable scale lengths ( $> 1000$  km) are necessary.  $P \rightarrow S$  conversion appears to be more efficient than  $P \rightarrow P$  scattering, but not efficient enough to explain the coda. The horizontal dashed line in figure 9 represents a 10% perturbation, which can be assumed to be a reasonable perturbation. This level is reached only for large scatterer size. For waves arriving at longer lapse times even stronger scatterers are needed, because the scattered wave has traveled longer distance and has

therefore smaller amplitude. Thus, single scattering of body waves is not a likely mechanism to explain the *P* or *PP* coda at these periods.

Apparently, unrealistically strong scatterers are needed to explain the coda level. It was shown by Dubendorff & Menke (1980), on the basis of model experiments, that single scattering theory tends to underestimate the scattering, and thus to overestimate the scatterers. Sato (1989) showed that single scattering theory breaks down at long lapse times. A comparison between single and multiple scattering theory was done by Frankel & Wennerberg (1987). They calculated coda amplitude in a random medium using an energy-conserving energy-flux-model and compared the coda amplitude decay with the decay predicted by the single-scattering model. The latter model predicts a decay that is too steep, and if used to estimate scatterer strength and size from observed coda decay rate, would overestimate these values. However, it is not likely that the overestimation is so large that multiple scattering would produce acceptable scatterers in the present problem. Even for the relatively short lapse time of 50 s (which is only 5 cycles at 0.1 Hz), to obtain reasonable scatterer strength or size, single scattering theory would have to overestimate the scatterers (either in strength or size) by a factor of more than 10, which is unlikely.

Therefore, it appears that body wave to body wave scattering is not an adequate explanation of coda generation at the low frequencies studied here. This is in agreement with the beamforming results. This does not hold for higher frequencies (generally higher than about 1 Hz), as is shown by various authors mentioned before. The next section deals with the coupling of surface and body waves and applies some linearized theories to realistic scatterers.

## 6. Rayleigh to *P* coupling.

The beamforming results indicate that the scattered waves in the *PP*- and *S*-wave coda are predominantly surface waves. These must have been excited by the earlier arriving body waves, as the signal-to-noise ratio is very good for most of the events. Prime candidate scatterers are surface topography and topography on the Moho; deeper heterogeneity will not contribute much to the scattering, as the Rayleigh waves at about 10 s do not penetrate very deep into the upper mantle. This explanation of body wave coda has been suggested earlier in the literature, for frequencies higher than 1 Hz (e.g. Key 1967; Greenfield 1971; Baumgardt 1985; Gupta et al. 1990). The coupling of Rayleigh to *P* has received much attention in experimental (e.g., McGarr & Alsop 1967; Gangi & Wesson 1978 and references therein) and numerical studies (e.g. Fuyuki & Matsumoto 1980; Boore et al. 1981; Hill & Levander 1984; Levander & Hill 1985; Stead & Helmberger 1988). From these studies it appears that reasonable inhomogeneity can explain scattered Rayleigh waves with amplitudes comparable to coda levels, observed here. It must be noted that all studies mentioned above are two-dimensional, so that Rayleigh wave amplitudes are not affected by geometrical spreading.

To investigate the efficiency of *P*-to-*R* coupling, theory on this subject was applied to realistic scatterers. The coupling due to topography on the Moho is treated by Odom (1986). This (linearized) theory is valid for a 2-D acoustic medium with small changes in the depth to the interface between an layer over a half space. The model is a dipping Moho: a dip of 10% over a distance of 200 km. This is a quite reasonable topography, it is common under western Europe (e.g., Meissner 1986; Luosto et al. 1990). Figure 10 shows the

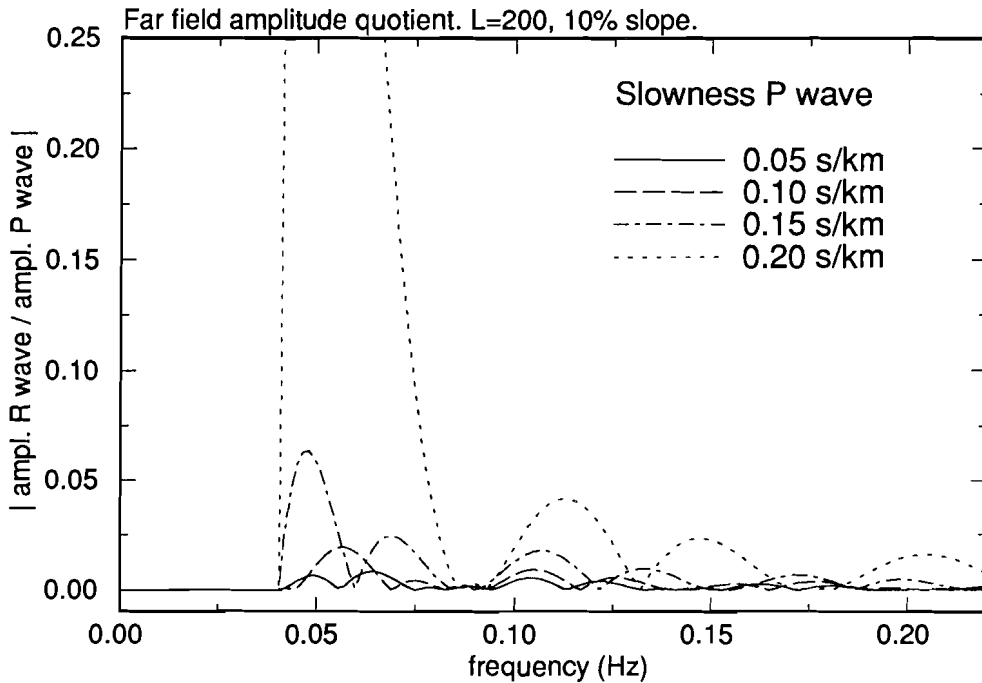


Figure 10. Amplitude of discrete mode excited by a  $P$  wave of amplitude 1, as a function of frequency and slowness of the  $P$  wave. The scatterer is a sloping Moho, the extent of the slope is 200 km, the slope is 10%. The medium is 2-D and acoustic.

amplitude of the converted guided mode (normalized by the amplitude of the plane incoming wave) as a function of both frequency and slowness of the direct wave. The results are independent of distance to the scatterer, as the direct wave is plane and the discrete mode is not affected by geometrical spreading. The figure shows that discrete modes with amplitudes of about 5% of that of the direct wave are possible, ignoring the very large-amplitude mode that is excited by a  $P$  wave traveling almost horizontally (slowness 0.2 s/km). At higher frequencies the coupling is less efficient; this is because the penetration depth becomes shallower. It should be noted that the coupling may be weaker in an elastic medium, where the phase velocity of  $P$  is quite different from that of the Rayleigh modes. Therefore, this model probably overestimates the  $P$ -to- $R$  coupling.

Gilbert & Knopoff (1960) described the  $P$ -to- $R$  coupling due to surface topography in two dimensions. The 2-D model is an elastic half space with a triangular hill, with 10%-sloping sides. The half space has  $\alpha = 6$ ,  $\beta = 3.5$  and  $c_R = 3.4$  km/s, where  $c_R$  is the Rayleigh wave velocity. Figure 11 shows the amplitude of the Rayleigh wave scattered from a plane  $P$  wave (with amplitude 1), as a function of slowness of the plane direct wave. For this scatterer a Rayleigh wave with an amplitude of 10% of that of the  $P$ -wave is predicted.

A result for a similar scatterer in three dimensions can be obtained with an extension to 3-D of Gilbert & Knopoff's theory by Hudson (1967) and Hudson & Boore (1980). The

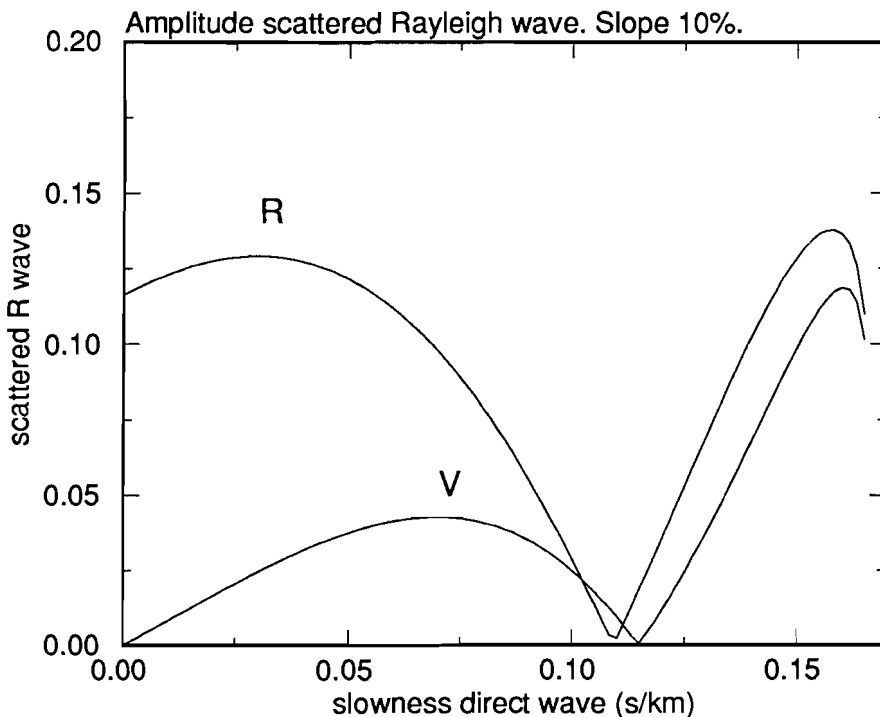


Figure 11. Amplitude of vertical and radial component of Rayleigh wave scattered from a  $P$  wave with amplitude 1, as a function of the slowness of the  $P$  wave. The scatterer is a 2-D triangular hill, with 10% sloping sides.

3-D scatterer is a Gaussian-shaped hill, with an area of  $20 \times 20 \text{ km}^2$  and a height of 1 km. The (plane)  $P$ -wave is vertically incident. The results are shown in figure 12. The results show that at a period of 10 s the coupling is most efficient. Due to geometrical spreading, the Rayleigh wave that has an amplitude (relative to  $P$ ) of about 10% very near the scatterer, quickly dies out to only 1% at 100 km distance (this is indicated by the two curves in the plot).

From these simple calculations and the results presented in the literature, it appears that the  $R$ -to- $P$  coupling in 2-D is quite efficient, especially at topography on the surface. Predicted Rayleigh wave amplitudes agree with those observed for realistic scatterers at frequencies of about 0.1 Hz. As there is no geometrical spreading, the coda thus created has a constant amplitude in time, assuming that there are enough scatterers and ignoring multiple scattering. The result for the three-dimensional Gaussian-shaped hill shows that the amplitude of the scattered waves quickly decays. However, a constant coda level is observed. The question is thus whether the increased number of scatterers and the increased strength of the scatterers (in 3-D a scatterer is measured by its volume instead of its area in 2-D) outweigh the stronger spreading in 3-D.

A simple argument shows that the increased number of scatterers indeed weighs up to the

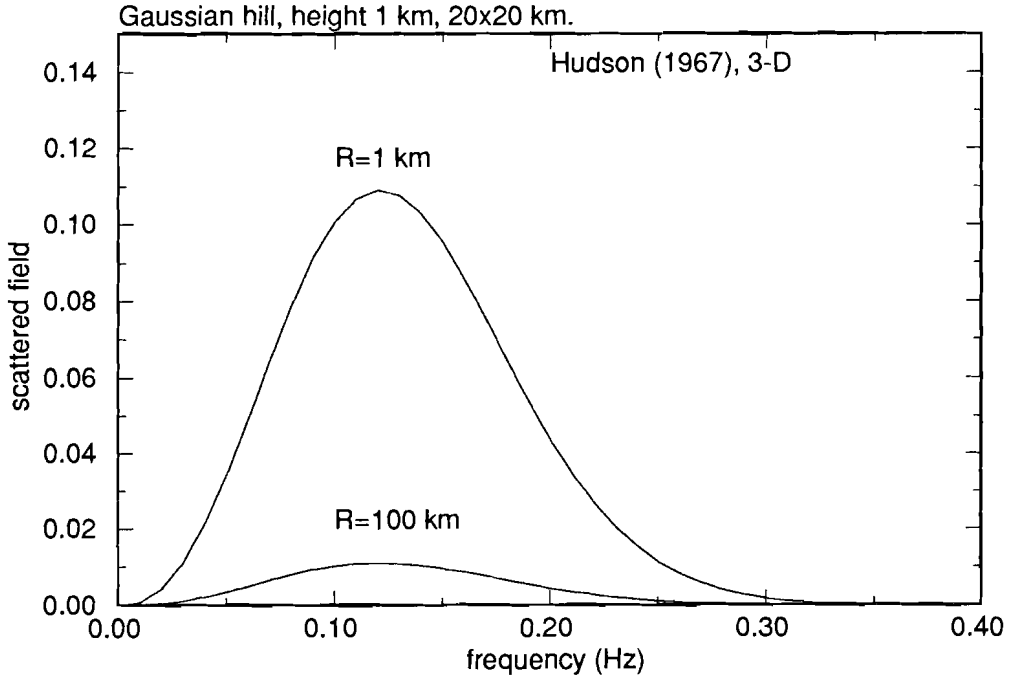


Figure 12. As figure 11, now for a 3-D gaussian-shaped hill with height 1 km. The area of the hill is about 20x20 km<sup>2</sup>. Due to geometrical spreading, the Rayleigh wave quickly decays.

stronger geometrical spreading. The geometry of the problem in 3-D is sketched in figure 13. The receiver R is located inside a scattering region, in which the scatterers are assumed to be distributed randomly. A plane body wave crosses the region, exciting Rayleigh waves. The scatterers that contribute to the coda between lapse times  $t$  and  $t + \Delta t$  are approximately located on a circular region  $E$  focused on the receiver;  $E$  is bounded by circles of radius  $r = c_R t$  and  $r + \Delta r = c_R(t + \Delta t)$ , with  $c_R$  the Rayleigh wave velocity. Assuming that this region is filled entirely with Gaussian-shaped hills with half-width  $L$ , the number of scatterers inside the region  $E$  is then approximately  $2r\Delta r/L^2$ . At the receiver the scattered waves can be written as

$$u = \sum_i \frac{A_i}{\sqrt{r_i}} e^{i\phi_i} \tag{2}$$

where the sum is over all scatterers in  $E$  and the distance  $r_i$  is the distance scatterer-receiver. It is assumed that the amplitude of the body wave is the same at scatterer and receiver, so it does not enter the formula. The power in the scattered wave is then

$$P = |u|^2 = \sum_i \sum_j \frac{A_i A_j^*}{\sqrt{r_i r_j}} e^{i(\phi_i - \phi_j)} \tag{3}$$

Assuming the waves add randomly (this is justified if  $\Delta t$  is longer than the duration of



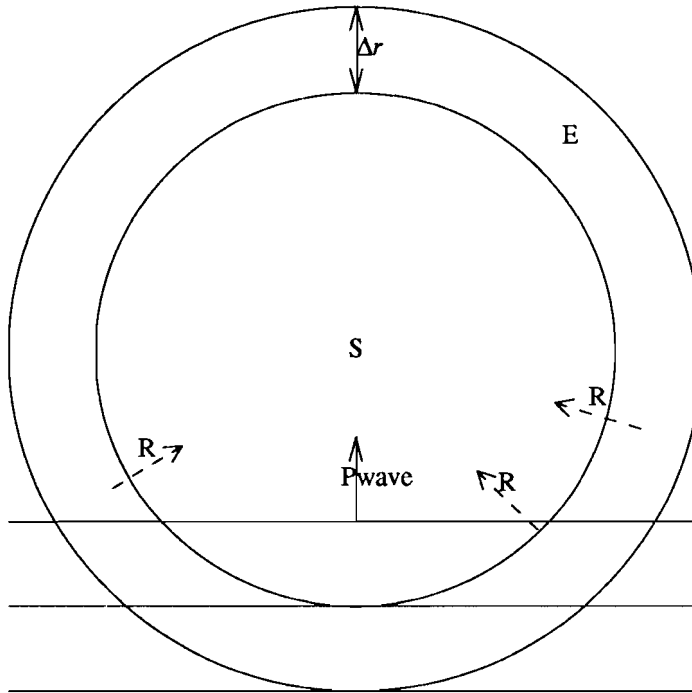


Figure 13. Geometry of the Rayleigh-to- $P$  coupling problem. A  $P$  wave crosses the region, exciting Rayleigh waves at homogeneously distributed scatterers. Scattered waves arriving between lapse times  $t$  and  $t + \Delta t$  are scattered by scatterers located on a surface, which can approximately be represented by the surface  $E$  bounded by two circles focused on the station  $S$ . The width of the region  $E$  is  $\Delta r = c_R \Delta t$ , where  $c_R$  is the Rayleigh wave velocity.

individual scattered surface wave trains) and that  $\Delta r \ll r$ , a major contribution in the sum comes whenever  $i = j$

$$P \approx \sum_i \frac{|A_i|^2}{r_i} = n \frac{|A_i|^2}{r_i} = \frac{2\Delta r |A|^2}{L^2} \quad (4)$$

The power is now independent of the distance from scatterer to receiver and thus independent of lapse time. Using this expression and the theory of Hudson (1967) for the Gaussian-shaped hill, a (constant) coda level of about 10-15% of that of the preceding body wave is found. This compares well with the value of about 20%, commonly observed for the events used here.

Thus, it is possible to explain the surface waves in the coda of  $PP$ - and  $S$ -waves as scattered from body waves. Surface topography alone can explain the greater part of the coda amplitude. It should be noted that at higher frequencies the coupling at surface topography is even more efficient. (Using his 3-D theory, Hudson (1967) explained a Rayleigh wave observed in the short-period (about 1 Hz) coda of  $P$  at the Eskdalemuir array (Key 1967) as excited by  $P$  at the slopes of a hill nearby the array.) Coupling at the Moho is probably weaker, but will increase the coda level. There are a number of assumptions in the above

reasoning that affect the result. First, there are other possibilities of coupling Rayleigh to *P*. Coupling by volume heterogeneities will increase the coda level. Unfortunately, there is no theory available treating volume scatterers, so their importance can not be estimated. Second, Hudson's theory is valid only for low-angle topography. Steep structures may cause a much larger coupling. Third, it is implicitly assumed that the Born approximation is valid for *P*-to-*R* scattering; this is supported by the spectra shown in figure 8. This is not the case for the surface waves at the periods studied here (about 10 s). Snieder (1988) showed that for periods shorter than about 20 s the surface waves are in the strong scattering domain. Therefore, for sufficiently long lapse times, the surface waves in the body wave coda observed at GRF have had multiple interactions with the inhomogeneities between the point of excitation and GRF. This will cause a certain amount of energy to be scattered out of the coda system, e.g. by reverse coupling to body waves. Not accounting for this energy loss leads to underestimation of scatterer strength and/or size. Some of the surface waves recorded in the coda may have been scattered by heterogeneities that are essentially two-dimensional. Such structures near GRF include the Alps, ocean-continent boundaries and the Rhine Graben. Rayleigh waves thus scattered have a geometrical spreading that is much less than  $1/\sqrt{r}$ ; 2-D experiments presented in the literature (e.g. Boore et al. (1981)) show that in this case it is not very difficult to obtain realistic Rayleigh wave amplitudes, even at long lapse times.

With the scattering argument the anomalously large *PP* coda amplitude of event 4 can be explained. This event lies near the caustic in the *PP* triplication range (distance  $58.7^\circ$ ) where *PP* waves arrive with large amplitude in a long time interval. Therefore, the *PP*-wave train is long and energetic. The *PP*-waves arrive consecutively (and thus do not add their amplitudes) and each of them excites a long coda of scattered surface waves. These coda do overlap, creating a large-amplitude coda.

## 7. Discussion.

The beamforming results show that the long-period coda of *PP*- and *S*-waves consist of scattered surface waves. 2-D experimental and numerical modeling results reported in the literature and the simple calculations presented in this study suggest that the coupling of Rayleigh to *P*- or *S*-waves is efficient enough to excite a coda level comparable in strength and duration to observations. The 3-D coupling coefficients predicted by the theory of Hudson (1967) for a smooth topography at the surface ( $20 \times 20 \text{ km}^2$  with a height of 1 km) agree well with observed coda amplitudes. When a dense, random coverage of such scatterers is assumed, a constant coda level can be predicted, which is close to observed values. Such a dense coverage is not as unrealistic as it may seem, as a quick glance at topographic maps shows. The scatterer used here is in fact very smooth when compared to actual topography. Surface topography is not the only mechanism to couple Rayleigh to *P*-waves, but 2-D calculations suggest that coupling due to variations in Moho depth is not as efficient. Unfortunately, no theory is available that includes the coupling due to volume heterogeneities (e.g. in the crust). The shortcoming of the theories of both Hudson (1967) and Odom (1986) are that they are valid for smooth topography only. There is also the possibility that the greater part of the observed scattered surface waves is actually scattered by obstacles that are essentially 2-D, as e.g. ocean-continent boundaries, mountain chains. Such waves have a much less severe geometrical spreading than waves scattered by a

---

feature that is finite (as compared to the wavelength) in both horizontal directions. It is not expected that body wave to body wave scattering contributes a significant amount to the coda level; in order to cause a scattered wave of reasonable amplitude either unacceptably large scattering volumes or unacceptably strong scatterers are needed. This is in agreement with the observation that very few scattered waves have a body wave slowness.

Explaining the *PP* and *S* coda as scattered surface waves raises the question why such waves are not observed in the long-period coda of *P*. For short periods scattered surface waves are observed after *P* (Key 1967; Langston 1989; Gupta et al. 1990). The explanation is given by the reflectivity synthetics shown in figures 6,7 and 8 of chapter 2. The synthetics show that the signal from the spherically symmetric Earth is rather strong between *P* and *PP*. There is hardly such signal after the *PP/PPP*-wave train, although the slowness interval that was selected contained slownesses up to that of the *S*-wave. Therefore, the coherent signal of the 1-D Earth may obscure scattered waves between *P* and *PP*, but after *PP* there is no such signal. It must also be noted that as nearly all events are within the *PP* triplication distance, the *PP* coda is caused by more than one *PP*-wave, whereas all scattered waves arriving before *PP* are scattered from a single *P*-wave. Furthermore, the coupling between Rayleigh and *P* is not as efficient as that between Rayleigh and *PP*. This can be seen in figure 10. Away from the nodes in the individual curves, the Rayleigh wave amplitude increases with increasing slowness at fixed frequency.

In contrast with the long-period coda of *P*, from which information about the layered structure of the upper mantle can be derived (chapter 2), the scattered surface waves carry only information about lateral heterogeneity of the crust and possibly the lithosphere, at very long periods. It is not feasible to extract deterministic characteristics of the scatterers from the coda, such as location and/or scale length and strength. This is due to the intrinsic complexity of scattered coda waves. At any moment in the coda interval scattered waves arrive from possibly all azimuths. More specifically, at any moment there is probably more than one wave arriving from the same azimuth. Such a complex wave field cannot be handled in a deterministic way; a statistical approach would be more fruitful. However, one obtains information about the crust only, and, judging from the efficiency of *R*-to-*P* coupling at surface topography, mainly about surface features. Of course there are more efficient ways to obtain such information.

## References.

- Aki, K., 1973. Scattering of *P* waves under the Montana Lasa, *J. Geophys. Res.*, **78**, 1335–1346.  
 Aki, K. & B. Chouet, 1975. Origin of coda waves: source, attenuation and scattering effects, *J. Geophys. Res.*, **80**, 3322–3342.  
 Aki, K. & P.G. Richards, 1980. *Quantitative Seismology*, Freeman, San Francisco.  
 Baag, C.-E. & C.A. Langston, 1985a. Shear-coupled *PL*, *Geophys. J. R. astron. Soc.*, **80**, 363–385.  
 Backus, G. & F. Gilbert, 1968. The resolving power of gross Earth data, *Geophys. J. R. astron. Soc.*, **16**, 169–205.  
 Baumgardt, D.R., 1985. Comparative analysis of teleseismic *P*-coda and *Lg* wave from underground nuclear explosions in Eurasia, *Bull. Seismol. Soc. Am.*, **75**, 1413–1433.  
 Boore, D.M., S.C. Harmsen & S.T. Harding, 1981. Wave scattering from a step change in surface topography, *Bull. Seismol. Soc. Am.*, **71**, 117–125.  
 Bostock, M.G. & B.L.N. Kennett, 1990. The effect of 3-D structure on *Lg* propagation patterns, *Geophys. J. Int.*, **101**, 355–365.

- Bouchon, M., 1982. The complete synthesis of seismic crustal phases at regional distances, *J. Geophys. Res.*, **87**, 1735-1742.
- Bungum, H. & J. Capon, 1974. Coda pattern and multipath propagation of Rayleigh waves at NOR-SAR, *Phys. Earth Planet. Int.*, **9**, 111-127.
- Capon, J., 1969. High resolution frequency-wavenumber spectrum analysis, *Proc. IEEE*, **57**, 1408-1418.
- Capon, J., 1970. Analysis of Rayleigh-wave multipath propagation at LASA, *Bull. Seismol. Soc. Am.*, **60**, 1701-1731.
- Dost, B., 1990. Upper mantle structure under western Europe from fundamental and higher mode surface waves using the NARS array, *Geophys. J. Int.*, **100**, 131-151.
- Dubendorff, B. & W. Menke, 1989. Time-domain apparent-attenuation operators for compressional and shear waves: experiment versus single-scattering theory, *J. Geophys. Res.*, **91**, 14023-14032.
- Flatté, S.M. & R.-S. Wu, 1988. Small-scale structure in the lithosphere and asthenosphere deduced from arrival time and amplitude fluctuations at NORSAR, *J. Geophys. Res.*, **93**, 6601-6614.
- Frankel, A. & Clayton, R.W., 1986. Finite difference simulations of seismic scattering: implications for the propagation of short-period seismic waves in the crust and models of crustal heterogeneity, *J. Geophys. Res.*, **94**, 6465-6489.
- Frankel, A. & L. Wennerberg, 1987. Energy-flux model of seismic coda: separation of scattering and intrinsic attenuation, *Bull. Seismol. Soc. Am.*, **77**, 1223-1251.
- Fuyuki, M. & Matsumoto, Y., 1980. Finite-difference analysis of Rayleigh wave scattering at a trench, *Bull. Seismol. Soc. Am.*, **70**, 2051-2070.
- Gangi, A.F. & R.L. Wesson, 1978. P-wave to Rayleigh-wave conversion. coefficients for wedge corners: model experiments, *J. Comp. Phys.*, **29**, 370-388.
- Gao, L.S., L.C. Lee, N.N. Biswas & K. Aki, 1983. Comparison of the effect between single and multiple scattering on coda waves for local earthquakes, *Bull. Seismol. Soc. Am.*, **73**, 377-389.
- Gilbert, F. & L. Knopoff, 1960. Seismic scattering from topographic irregularities, *J. Geophys. Res.*, **65**, 3437-3444.
- Greenfield, R.J., 1971. Short-period P-wave generation by Rayleigh-wave scattering at Novaya Zemlaya, *J. Geophys. Res.*, **76**, 7988-8002.
- Gupta, I.N., C.S. Lynnes & R.A. Wagner, 1990. Broadband f-k analysis of array data to identify sources of local scattering, *Geophys. Res. Lett.*, **17**, 183-186.
- Helmberger, D.V. & G.R. Engen, 1980. Modeling the long-period body waves from shallow earthquakes at regional ranges, *Bull. Seismol. Soc. Am.*, **70**, 1699-1714.
- Hill, N.R. & A.R. Levander, 1984. Resonances of low-velocity layers with lateral variations, *Bull. Seismol. Soc. Am.*, **74**, 521-537.
- Högbom, J.A., 1974. Aperture synthesis with a non-regular distribution of interferometer baselines, *Astron. Astrophys. Suppl.*, **15**, 417-426.
- Hudson, J.A., 1967. Scattered surface waves from a surface obstacle, *Geophys. J. R. astron. Soc.*, **13**, 441-458.
- Hudson, J.A. & D.M. Boore, 1980. Comments on "Scattered surface waves from a surface obstacle" by J.A. Hudson, *Geophys. J. R. astron. Soc.*, **60**, 123-127.
- Kennett, B.L.N., 1987. Observational and theoretical constraints on crustal and upper mantle heterogeneity, *Phys. Earth Planet. Int.*, **47**, 319-332.
- Key, F.A., 1967. A signal-generated noise recorded at the Eskdalemuir Seismometer Array Station, *Bull. Seismol. Soc. Am.*, **57**, 27-38.
- King, D.W., R.A.W. Haddon & E.S. Husebye, 1975. Precursors to PP, *Phys. Earth Planet. Int.*, **10**, 103-127.
- Korn, M., 1988. P-wave coda analysis of short-period array data and the scattering and absorptive properties of the lithosphere, *Geophys. J.*, **93**, 437-449.
- Langston, C.A., 1989. Scattering of teleseismic body waves under Pasadena, California, *J. Geophys. Res.*, **94**, 1935-1951.

- Levander, A.R. & N.R. Hill, 1985. *P-SV resonances in irregular low-velocity surface layers*, *Bull. Seismol. Soc. Am.*, **75**, 847–864.
- Levshin, A. & K.-A. Berteussen, 1979. Anomalous propagation of surface waves in the Barents Sea as inferred from NORSAR recordings, *Geophys. J. R. astron. Soc.*, **56**, 97–118.
- Luosto, U., T. Tiira, et al., 1990. Crust and upper mantle structure along the DSS profile in SE Finland, *Geophys. J.*, **101**, 89–110.
- McGarr, A. & L.E. Alsop, 1967. Transmission and reflection of Rayleigh waves at vertical boundaries, *J. Geophys. Res.*, **72**, 2169–2180.
- Meissner, R., 1986. *The continental crust, a geophysical approach*, Academic Press, Orlando.
- Odom, R., 1986. A coupled mode examination of irregular waveguides including the continuum spectrum, *Geophys. J. R. astron. Soc.*, **86**, 425–453.
- Palmer, D.R., M.G. Brown, F.D. Tappert & H.F. Bezdek, 1988. Classical chaos in nonseparable wave propagation problems, *Geophys. Res. Lett.*, **15**, 569–572.
- Sato, H., 1984. Attenuation and envelope formation of three-component seismograms of small local earthquakes in randomly inhomogeneous lithosphere, *J. Geophys. Res.*, **89**, 1221–1241.
- Sato, H., 1989. Is the single scattering model invalid for the coda excitation at long lapse time?, *Pure Appl. Geophys.*, **128**, 43–47.
- Sereno, T.J. & Orcutt, J.A., 1988. Synthetic  $P_n$  and  $S_n$  phases and the frequency dependence of  $Q$  of oceanic lithosphere, *J. Geophys. Res.*, **92**, 3541–3566.
- Snieder, R., 1988. Large scale waveform inversions of surface waves for lateral heterogeneity - II: Application to surface waves in Europe and the Mediterranean, *J. Geophys. Res.*, **93**, 12067–12080.
- Spakman, W. & G. Nolet, 1988. Imaging algorithms, accuracy and resolution in delay time tomography, in: *Mathematical Geophysics*, edited by N.J. Vlaar, G. Nolet, M.J.R. Wortel and S.A.P.L. Cloetingh, p 155–187, Reidel, Dordrecht, The Netherlands.
- Stead, R.J. & D.V. Helmberger, 1988. Numerical-analytical interfacing in two dimensions with applications to modeling NTS seismograms, *Pure Appl. Geophys.*, **128**, 157–193.
- Toksöz, M.N., A.M. Dainty & E.E. Charrette, 1989. Coherency of ground motion at regional distances and scattering, *Phys. Earth Planet. Int.*, submitted.
- Wu, R. & K. Aki, 1985b. Elastic wave scattering by a random medium and the small-scale inhomogeneities in the lithosphere, *J. Geophys. Res.*, **90**, 10261–10273.

### Appendix A: beamforming.

Suppose we have  $K$  stations with recorded time series  $s_j(t)$  and Fourier transforms of these traces  $S_j(\omega)$ . We wish to find the spectrum of slownesses present in the wave field. The relation between the data  $S_j(\omega)$  and the slowness spectrum is

$$S_j(\omega) = \int_{-\infty}^{\infty} P(\mathbf{p}) e^{-i\omega \mathbf{p} \cdot \mathbf{r}_j} d^2 p \quad (\text{A1})$$

where  $P(\mathbf{p})$  is the true slowness spectrum,  $\mathbf{p}$  is the horizontal slowness vector and  $\mathbf{r}_j$  is the vector pointing from a reference point (e.g. the centre of the array) to station  $j$ . The factor  $\exp(-i\omega \mathbf{p} \cdot \mathbf{r}_j)$  is the phase shift, relative to the reference point, of a wave with slowness  $\mathbf{p}$  due to the distance  $\mathbf{r}_j$  projected on the direction of the slowness  $\mathbf{p}$ . The integration limits are in practice not infinity; a suitable upper limit may be chosen, based on physically expected minimum velocities. The simplest estimate  $\hat{P}$  of the true slowness spectrum  $P(\omega, \mathbf{p})$  at frequency  $\omega$  is just a sum of the data  $S_j(\omega)$ , each delayed with a phase delay of  $\mathbf{p} \cdot \mathbf{r}_j$

$$\hat{P}(\omega, \mathbf{p}) = \sum_{j=1}^K w_j S_j(\omega) e^{i\omega \mathbf{p} \cdot \mathbf{r}_j} \quad (\text{A2})$$

Weights  $w_j$  can be used to control the slowness resolution; in the conventional beamforming technique these are set to 1. The components  $S_j$  are normalized to unit length to avoid artifacts due to uneven station calibration or amplitude variations across the array. Dispersive waves may also be dealt with by averaging the energy in the spectrum over a frequency interval

$$\int_{\omega_1}^{\omega_2} |\hat{P}(\omega, \mathbf{p})|^2 d\omega \quad (\text{A3})$$

The array response is obtained by inserting a monochromatic wave in (A2). For a wave with slowness  $\mathbf{p}_o$ , we have  $S_j = \exp(-i\omega \mathbf{p} \cdot \mathbf{r}_j)$  and the array response is then simply

$$R(\mathbf{p}, \mathbf{p}_o) = \sum_{j=1}^K w_j e^{i\omega(\mathbf{p}-\mathbf{p}_o) \cdot \mathbf{r}_j} \quad (\text{A4})$$

The array response obtained with the conventional method for a frequency of 0.05 Hz and a vertically incident wave is shown in figure 14<sup>a</sup>. The shape of the array is reflected in its array response: elongated in the E-W direction, the direction in which the array is narrowest. The presence of side lobes and the finiteness of the main lobe are the result of the finite sampling in space of the wave field.

The resolving power of this method is poor, so there is need for an optimization scheme. Equation (A1) can be viewed as a linear inverse problem of the form  $\mathbf{d} = \mathbf{G}\mathbf{m}$  (dropping integrals), where  $\mathbf{G}$  is some operator (in this case  $\exp(-i\omega \mathbf{p} \cdot \mathbf{r}_j)$ ) acting on the real model  $\mathbf{m}$  (the slowness field) to yield the data vector  $\mathbf{d}$  ( $S_j(\omega)$ , the phase of frequency  $\omega$  at station  $j$ ). The problem is now to find the best possible estimate  $\hat{\mathbf{m}}$  of the real model  $\mathbf{m}$ . The theory of Backus & Gilbert (1968) can be applied here.

The data  $\mathbf{d}$  are related to the true spectrum  $P(\mathbf{p})$  through an operator  $\mathbf{G}(\mathbf{p})$  (replacing the integration limits by an estimated maximum slowness  $p_{\max}$ )

$$\mathbf{d} = \int \int_{|\mathbf{p}| \leq p_{\max}} \mathbf{G}(\mathbf{p}) P(\mathbf{p}) d^2 p \quad (\text{A5})$$

with  $\mathbf{G}$ ,  $P$  and  $\mathbf{d}$  as described above. The estimate  $\hat{P}(\mathbf{p})$  is linearly related to the data

$$\hat{P}(\mathbf{p}_o) = \sum_{j=1}^K a_j(\mathbf{p}_o) d_j = \tilde{\mathbf{a}} \mathbf{d} \quad (\text{A6})$$

Due to the finite number of data, this estimate is a blurred image of the real spectrum. This can be written as

$$\hat{P}(\mathbf{p}_o) = \int \int_{|\mathbf{p}| \leq p_{\max}} A(\mathbf{p}_o, \mathbf{p}) P(\mathbf{p}) d^2 p \quad (\text{A7})$$

where  $A$  is the averaging kernel that blurs the real spectrum. For the conventional method,  $A$  is just the array response  $R$  given by (A4). Putting (A5) into (A6) and using (A7) we

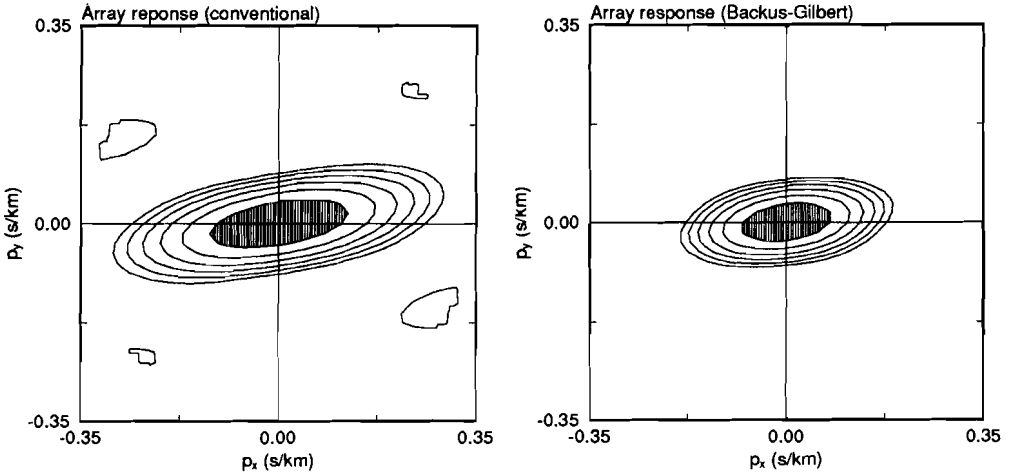


Figure 14. Array response of GRF for a frequency of 0.05 Hz. a) response with the conventional beam forming method. Two side lobes can just be seen at the edges of the spectrum. b) response obtained with the Backus-Gilbert inversion. The main lobe is smaller and the side lobes are removed. The line contours are in steps of 2 dB down from the maximum.

find an expression for the averaging kernel  $A$

$$A(\mathbf{p}_0, \mathbf{p}) = \sum_{j=1}^K a_j(\mathbf{p}_0) G_j(\mathbf{p}) = \tilde{\mathbf{a}}(\mathbf{p}_0) \mathbf{G}(\mathbf{p}) \quad (\text{A8})$$

We now have to find those weights  $\mathbf{a}$  that give the optimum estimate  $\hat{P}$ . The most desirable kernel  $A$  would be a delta function, for then  $\hat{P}(\mathbf{p}_o)$  is closest to  $P(\mathbf{p}_o)$  in (A7). Therefore, we would want a kernel  $A$  that is as close to a delta function as possible; this kernel is found by minimizing

$$J = \int \int_{|\mathbf{p}| \leq p_{\max}} \left( A(\mathbf{p}, \mathbf{p}_o) - \delta(\mathbf{p} - \mathbf{p}_o) \right)^2 d^2 p + \epsilon |\mathbf{a}|^2 \quad (\text{A9})$$

The second term in (A9) is added to ensure stability of the inverse in the presence of errors. One can see the parameter  $\epsilon$  as a trade-off between resolution (deltaness of the array response) versus variance (error in the result). The optimum value of  $\epsilon$  can be estimated by calculating the variance and resolution as a function of  $\epsilon$  and frequency; see Backus & Gilbert (1968) or Aki & Richards (1980, chapter 12).

Minimizing (A9) with respect to  $a_j$  and using (A8) gives the the solution for the optimum weights

$$\mathbf{a}_i(\mathbf{p}) = \sum_j (\mathbf{G}\tilde{\mathbf{G}} + \epsilon \mathbf{I})_{ij}^{-1} G_j(\mathbf{p}) \quad (\text{A10})$$

The matrix  $\mathbf{G}\tilde{\mathbf{G}}$  is defined as

$$\begin{aligned}
 (\mathbf{G}\tilde{\mathbf{G}})_{ij} &= \int \int_{|\mathbf{p}| \leq p_{\max}} G_i(\mathbf{p}) \tilde{G}_j(\mathbf{p}) d^2 p \\
 &= \int \int_{|\mathbf{p}| \leq p_{\max}} e^{i\omega \mathbf{p} \cdot (\mathbf{r}_i - \mathbf{r}_j)} d^2 p = \frac{2\pi p_{\max}}{\omega r_{ij}} J_1(\omega p_{\max} r_{ij})
 \end{aligned} \tag{A11}$$

where  $r_{ij}$  is the distance between stations  $i$  and  $j$ , and the integration is performed over the disc in  $\mathbf{p}$ -space with  $|\mathbf{p}| \leq p_{\max}$ . The expressions for  $A$  and  $\hat{P}$  now become

$$A(\mathbf{p}_o, \mathbf{p}) = \tilde{\mathbf{G}}(\mathbf{p}_o) (\mathbf{G}\tilde{\mathbf{G}} + \varepsilon \mathbf{I})^{-1} \mathbf{G}(\mathbf{p}) \tag{A12}$$

$$\hat{P}(\mathbf{p}_o) = \tilde{\mathbf{G}}(\mathbf{p}_o) (\mathbf{G}\tilde{\mathbf{G}} + \varepsilon \mathbf{I})^{-1} \mathbf{d} \tag{A13}$$

The array response for a vertically incident wave with frequency 0.05 Hz is shown in figure 14<sup>b</sup>. Working best at low frequencies, the Backus-Gilbert inversion has increased the resolution drastically. The main lobe is much smaller and side lobes are suppressed. At higher frequencies the side lobes are not removed totally from the spectrum, but as the position of these lobes is known, they can be removed from the spectrum with the CLEAN method (Högbon 1974).



## Chapter 4

### Topography of the 400 km discontinuity from observations of long-period $P_{400}P$ phases

**Summary.** New evidence is presented for lateral variation of the depth of upper mantle discontinuities. This evidence is derived from an analysis of amplitude of  $P_{400}P$  phases (underside reflections from the 400-km discontinuity) for great circle paths with bounce points in the upper mantle near the Hawaii-Emperor hot spot chain. The data are long-period, from WWSSN, SRO and RSTN stations. The epicentral distance ranges from  $95^\circ$  to  $115^\circ$ , which ensures that no other major phases arrive in the time window of interest. Observed  $P_{400}P$  amplitudes are compared with synthetic amplitudes for both a laterally homogeneous model and models including 3-D topography on the discontinuities. The observations permit models of topography that can be characterized by undulations of the 400-km discontinuity with wavelengths in the range 600-2000 km, and peak-to-peak amplitudes of the order of 30-40 km. Such long-wavelength structure causes quite strong (de)focusing of  $P_dP$  phases; synthetic  $P_dP$  wave forms reveal that topography not only affects its amplitude, but also smears out the  $P_dP$  wave, causing overlap between different  $P_dP$  phases. This mechanism is proposed as the reason for the lack of well-developed  $P_dP$  phases reported in the literature. The fact that the wave forms can be seriously distorted prohibits the use of  $P_dP$  phases in a wave form inversion. The curvature of the discontinuity can be explained as the signature of convection currents crossing the discontinuity. Although the data cover the Hawaii hot spot, no signature of a plume structure can be detected in

---

This chapter has been published as:

Neele, F. and R. Snieder, Topography of the 400 km discontinuity from observations of long-period  $P_{400}P$  phases, *Geophys. J. Int.*, 109, 670-682, 1991.

---

the data, probably because the resolving power of the data is too low.

## 1. Introduction

A detailed knowledge of the upper mantle discontinuities is of major importance to the understanding of the dynamical behaviour of the Earth. Lateral variations of the properties of the discontinuities may show the effect of convective processes in the upper mantle. Imaging such variations can put constraints on possible modes of convection. Seismic phases interacting with the discontinuity provide a direct means of measuring the properties of the discontinuity. These phases include conversions, either  $P$ - $S$  or  $S$ - $P$  (Barley, Hudson & Douglas 1982; Paulssen 1985, 1988; Kind & Vinnik 1988); reflections from the top of the discontinuities, producing  $P$ -related arrivals (Husebye & Madariaga 1970; Gutowski & Kanasewich 1974); and bottom-side reflections, producing either precursors to  $P'P'$  (Whitcomb & Anderson 1970; Haddon, Husebye & King 1977; Husebye, Haddon & King 1977; Nakanishi 1986, 1988; Davis, Kind & Sacks 1989) or to  $PP$  (Bolt 1970; Wajeman 1988). The results from these studies do not show a consistent picture, as far as depth and velocity contrast of the discontinuities are concerned. A large scatter in the reported depths of the upper mantle discontinuities appears. Furthermore, these discontinuity-related phases are observed only intermittently, suggesting that the reflectivity of the gradient zones varies laterally. Evidence for a locally pronounced 670-km transition was reported by Paulssen (1988) and Kind & Vinnik (1988). Richards (1972) showed that if the thickness of the velocity gradient exceeds a quarter of the seismic wavelength the amplitude of reflected phases significantly decreases. However, phases converted at the 670-km discontinuities have been observed on short-period records, suggesting a narrow transition. Another possible explanation for the scatter in depth and contrast of the upper mantle discontinuities is topography on the interfaces, causing focusing and defocusing (Richards 1972). Davis et al. (1989) used gaussian beam synthetics to show that even slight curvature of a discontinuity can cause significant (de)focusing of bottom-side reflected phases. Revenaugh & Jordan (1989) analysed travel time variations of  $ScS$ -related reverberations and reported topography on the 400 and 670 km discontinuity with wavelengths of the order of 500-1500 km and amplitudes of about 20 km. Detailed analysis of  $S$ -to- $P$  conversion at 670 km near subduction zones reveals conversion depths varying about 50 km (Barley et al. 1982; Bock & Ha 1984).

In the present study further evidence is presented for topography on upper mantle discontinuities.  $P_{400}P$  phases (precursors to  $PP$ ) are used to probe the 400 km discontinuity; topography causes amplitude variations as a function of the position of the bounce point. Data used are long-period seismograms from the WWSSN, RSTN and SRO stations. This restricts the resolution of the data, but it is shown that wavelengths smaller than the size of the Fresnel zone can be detected. The data sample the upper mantle in a region around the Hawaii-Emperor chain, providing an opportunity to look for a hot spot signature in the upper mantle.

## 2. Data

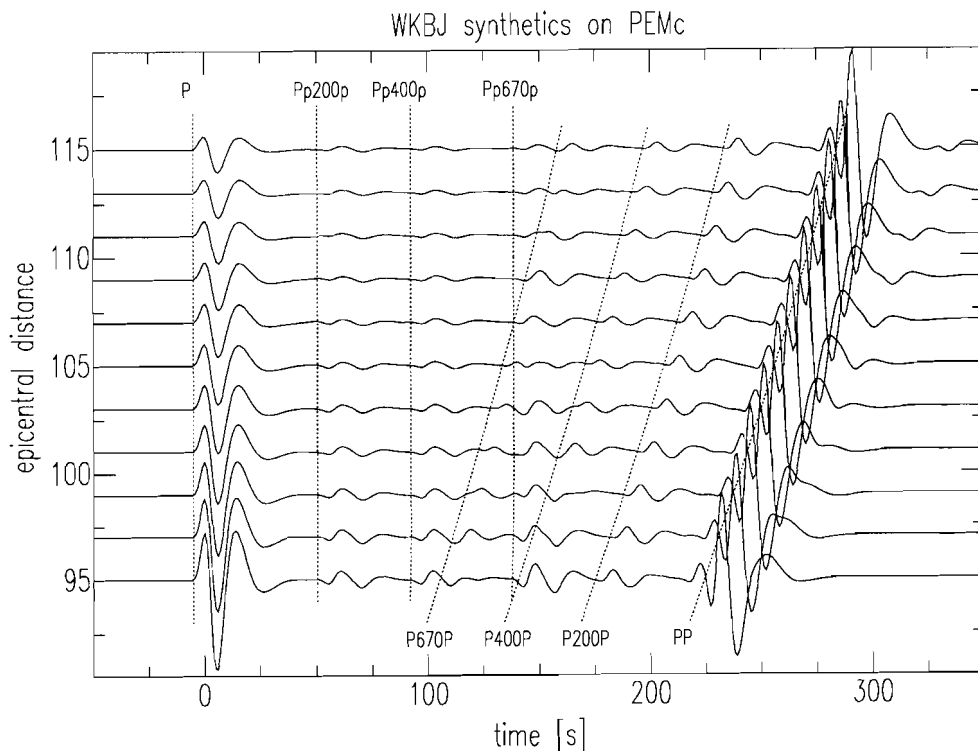
The data used in this study are long-period  $P_dP$  phases recorded at WWSSN and RSTN stations. Such phases are reflected at the underside of discontinuities in the upper mantle beneath the bounce point of  $PP$ , and arrive before  $PP$ . Data are taken from the NEIC data

Table 1. Events used in this study.

Date	Time	lat	lon	depth	$m_b$	station(s)
1980-06-25	23:18:20.40	-5.233	151.686	49.0	6.2	anno
1980-07-29	03:11:56.30	-13.101	166.338	48.0	5.9	anno
1980-09-26	15:20:37.10	-3.225	142.237	33.0	5.9	anno
1981-04-24	21:50:06.00	-13.426	166.421	33.0	6.1	anno
1981-10-25	03:22:15.57	18.048	-102.084	33.0	6.2	gumo
1981-11-06	16:47:49.10	-3.558	143.790	33.0	6.2	anno lon
1982-01-11	06:10:06.49	13.752	124.358	45.7	6.0	jas
1982-05-02	11:19:38.01	-29.318	-177.151	25.4	6.0	col
1982-06-07	06:52:37.37	16.607	-98.149	40.5	6.0	gumo
1982-08-05	20:32:52.95	-12.597	165.931	30.7	6.2	anno
1982-12-19	17:43:54.85	-24.133	-175.864	33.0	5.9	rsnt
1983-07-12	15:10:03.40	61.031	-147.286	37.0	6.1	snzo
1983-10-15	10:56:50.80	-8.101	156.311	7.4	5.9	anno rsnt rssid rson
1983-10-28	14:06:06.61	44.058	-113.857	10.0	6.2	ctao
1984-05-26	22:42:48.01	-10.889	164.119	33.0	5.7	anno rsnt rssid rson
1984-07-05	05:21:48.52	-6.085	154.423	29.7	6.1	anno rsnt rson
1984-08-31	15:42:12.18	-17.940	172.154	33.0	6.1	rsnt rssid rson
1984-09-10	03:14:08.96	40.346	-126.859	10.0	6.2	ctao
1984-09-28	00:03:35.46	-25.795	-176.014	26.8	6.4	rsnt
1985-07-03	15:55:49.00	-17.251	167.839	28.4	5.8	anno rsnt rssid rson
1985-09-26	07:27:48.98	-34.627	-178.694	33.0	6.5	col rsnt
1985-09-27	03:39:08.27	-9.805	159.844	29.6	6.3	anno rsnt rson
1985-10-05	15:24:02.25	62.257	-124.312	10.0	6.5	ctao
1985-12-21	01:13:21.09	-14.035	166.508	33.0	6.1	anno rsnt rson
1985-12-23	05:16:03.91	62.207	-124.302	10.0	6.4	ctao
1986-04-14	00:25:12.46	-13.946	166.977	33.0	5.8	rssid
1986-10-14	16:53:07.45	-4.985	153.584	33.0	6.1	anno rsnt rssid rson

set, available on CD-ROM. The events are selected to have the PP bounce point between event and station near the Hawaii-Emperor chain; the area selected is longitude 165°E to 150°W and latitude 10°N to 45°N. The epicentral distance range is constrained by the requirement that the  $P_dP$  phases arrive in a window uncontaminated by other phases. This restricts the epicentral distance to the interval 95-115°: at shorter ranges the  $Pp_{670}$   $p$  phase (a  $P$ -multiple between the surface and the 670-km discontinuity) arrives close to  $PP$ ; at longer ranges  $PKP$  enters the  $PP$  window. Only shallow events ( $h < 50$  km) are considered, to avoid  $pP$  or  $sP$  interfering with  $P_dP$  phases. As the  $PP$ -precursors have small amplitudes, the data must have a high signal-to-noise ratio. Finally, the source wavelet should be simple; data with a complex, long  $P$  wavelet were discarded. Scanning the NEIC data set selecting for epicentral distance, mid point location and source depth yielded 488 event-station pairs; of these only 49 met the remaining criteria (S/N, simple source pulse). Table 1 lists the event-station combinations used. Figure 6 shows the location of the bounce points of  $PP$  and  $P_dP$  phases near the Hawaii-Emperor chain. The dotted line gives the approximate location of the island chain. The azimuth of the great circle through each bounce point is indicated by a short line. Most great circles have an azimuth trending roughly SW-NE.

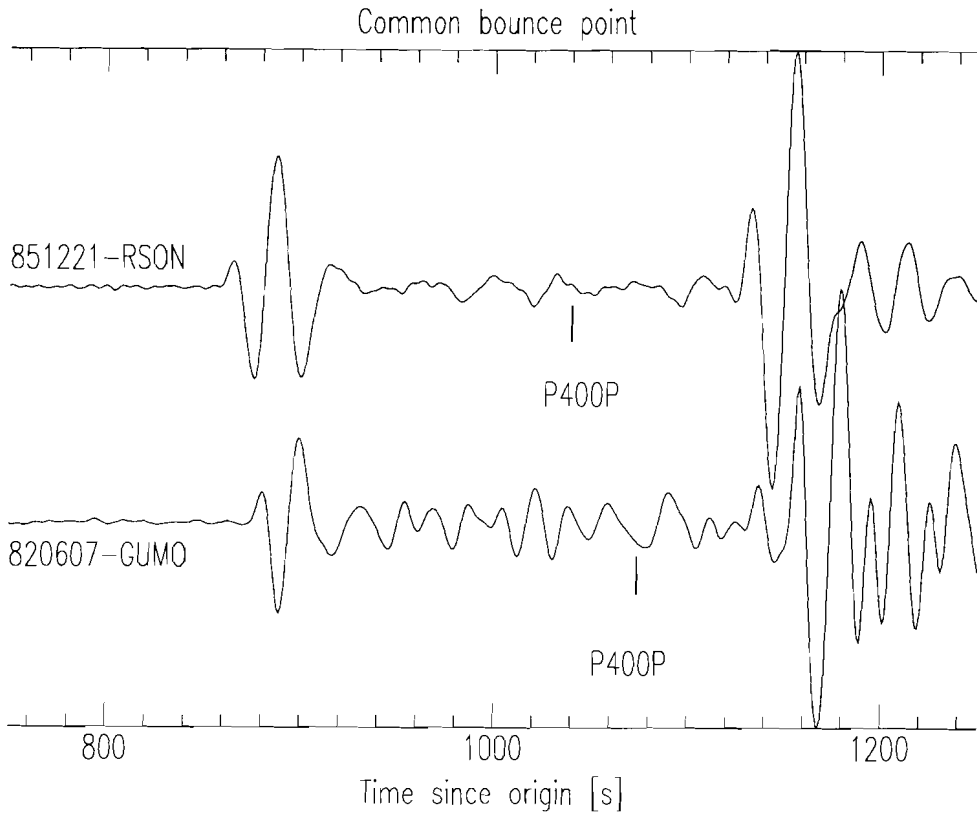
Figure 1 shows the response of a perfectly spherically symmetric Earth, between the  $P$  and  $PP$  phases. The synthetic seismograms are obtained with the WKBJ method, for an explosive source on PEMc (Dziewonski, Hales & Lapwood 1975), recorded on the vertical



*Figure 1.* Time-distance plot of synthetic seismograms obtained with the WKBJ method; the Earth model used is PEMc. *P*- and *PP*-related phases stand out clearly.

component at ranges from  $95^\circ$  to  $115^\circ$ . The synthetics are convolved with the instrument response of SRO station ANMO, and low-passed to have the same band-width as the data, i.e. periods longer than about 7 s. In the absence of noise and scattering,  $Pp_d p$  and  $P_d P$  phases can be seen as more or less impulsive arrivals across the plot, behaving coherently. The amplitude of the  $P_d P$  phases decreases slowly with increasing epicentral distance. For distances smaller than  $100^\circ$ ,  $Pp_{670} p$  interferes with  $P_{400} P$ . For the whole interval shown here,  $P_{670} P$  arrives in the same time interval as  $Pp_{670} p$  and, for shorter ranges,  $Pp_{400} p$ . Converted phases have not been included in the synthetics, as these have small amplitudes.

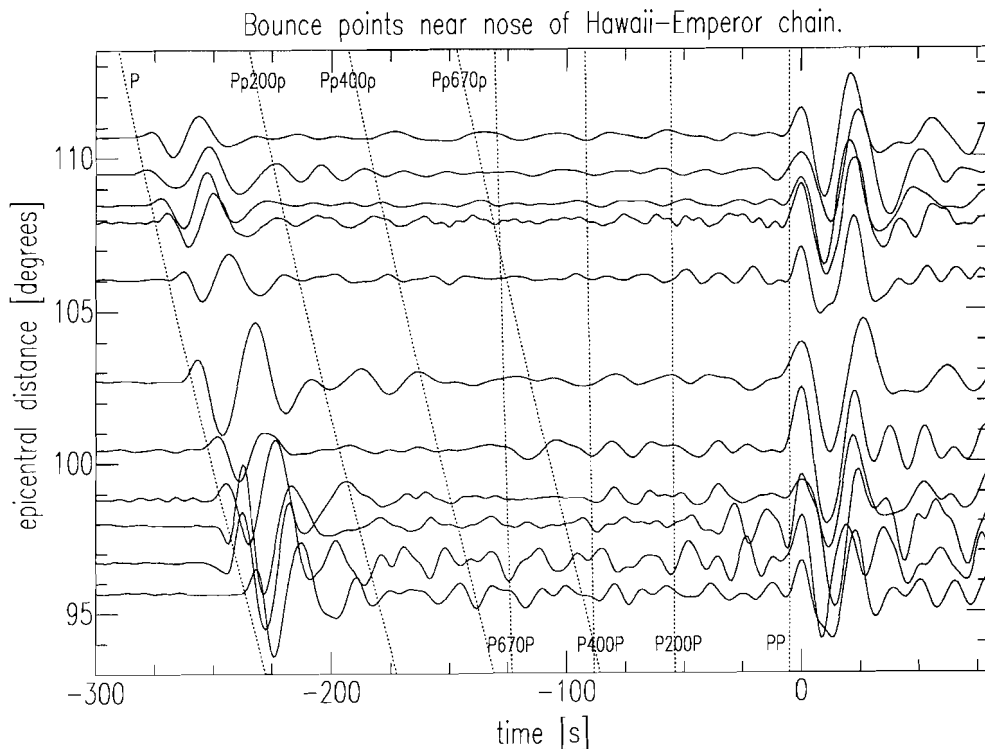
An example of the response of the real Earth in the same time interval in the seismogram is given in figure 2. The figure shows data for two great circle paths, for which the bounce point is almost the same; the azimuth of the great circle at the bounce point differs about  $45^\circ$ . The epicentral distances for the two records are  $107.0^\circ$  and  $110.9^\circ$ . Whereas the upper trace shows low amplitudes before *PP*, the lower trace contains much more energy in the same interval. The difference in epicentral distance can not explain the difference in energy content: one would expect the higher amplitudes at shorter distances rather than reversed (see figure 1). Both source wavelets are simple and short, eliminating a strong source effect. Note that the noise level is much lower than the signal between *P* and *PP*, therefore all energy is related to the Earth response; incoherence can not be attributed to noise. Neele



*Figure 2.* Two records of great circle paths with nearly the same bounce point for the *PP* (and *PP*-related) phases. The data shown are raw data; the traces have the same frequency content. Epicentral distances are  $106.9^\circ$  for 851221-RSON and  $110.9^\circ$  for 820607-GUMO. The arrival times for the  $P_{400}P$  phase for PEMc are indicated. The data show a striking difference in energy content of the *PP*-precursors.

& Snieder (1991) have shown from a slowness-azimuth analysis of data from the GRF-array that although the long-period coda of *PP* and *S* phases contain scattered waves, the coda of *P* (this includes all phases between *P* and *PP*) does not show significant deviations from great circle azimuth. Although the structure of the crust and lithosphere (where most of the long-period *P* coda is generated) beneath GRF will be different from that beneath the stations used here, it is assumed that the results from GRF apply here and that scattered waves can be ignored. As GRF is situated in a heterogeneous region, this is a reasonable assumption. The different great circle azimuth at the bounce point seems the only likely reason for the different behaviour of the *PP* precursors. This suggests that the properties of the discontinuities vary laterally, leading to an azimuth-dependent reflection coefficient.

Another example of the character of the data is given in figure 3, which is a time-distance plot of data for bounce points near the nose of the volcanic chain. The traces are lined up at the *PP* wave; any *PP*-related phases should line up with *PP* in such a plot. The arrival times of the main *P*- and *PP*-related phases (as predicted by the reference model PEMc)



*Figure 3.* A time-distance plot of some of the data, with  $PP$  bounce points near Hawaii. The traces are lined up at the  $PP$  wave arrival. Arrival times of some major  $PP$ - and  $P$ -related phases, as predicted by model PEMc, are indicated.

are indicated; these include top and bottom-side reflections at the upper mantle discontinuities. Although there seem to be some arrivals that can be traced from one record to the next, the first impression is that of disordered behaviour: no  $P_dP$  or  $Pp_d p$  phase stands out clearly. Comparing the synthetics in figure 1 with the data in figures 2 and 3 shows that the amplitude of the  $PP$  precursors relative to  $PP$  agree. The character of the phases between  $P$  and  $PP$  is totally different: there is little coherency among the records and no distinct arrivals can be identified.

From figures 1, 2 and 3 some characteristics of the data can be inferred, that are notably different than those of the synthetic seismograms. Firstly, the data do not show coherent arrivals. No  $P_dP$  (or  $Pp_d p$ ) phase can be followed across the entire epicentral distance range. In contrast to the synthetics, that show clear, impulsive arrivals between  $P$  and  $PP$ , the data have a more reverberatory character, in which no definite arrival times can be picked. Secondly, the data show a strong azimuth dependence. Any model describing  $P_dP$  phases must explain these characteristics, and they will be used in the next sections to select acceptable models.

The choice of long-period data has consequences for the resolution. For the employed periods (longer than about 7s), the upper mantle discontinuities are effectively step changes

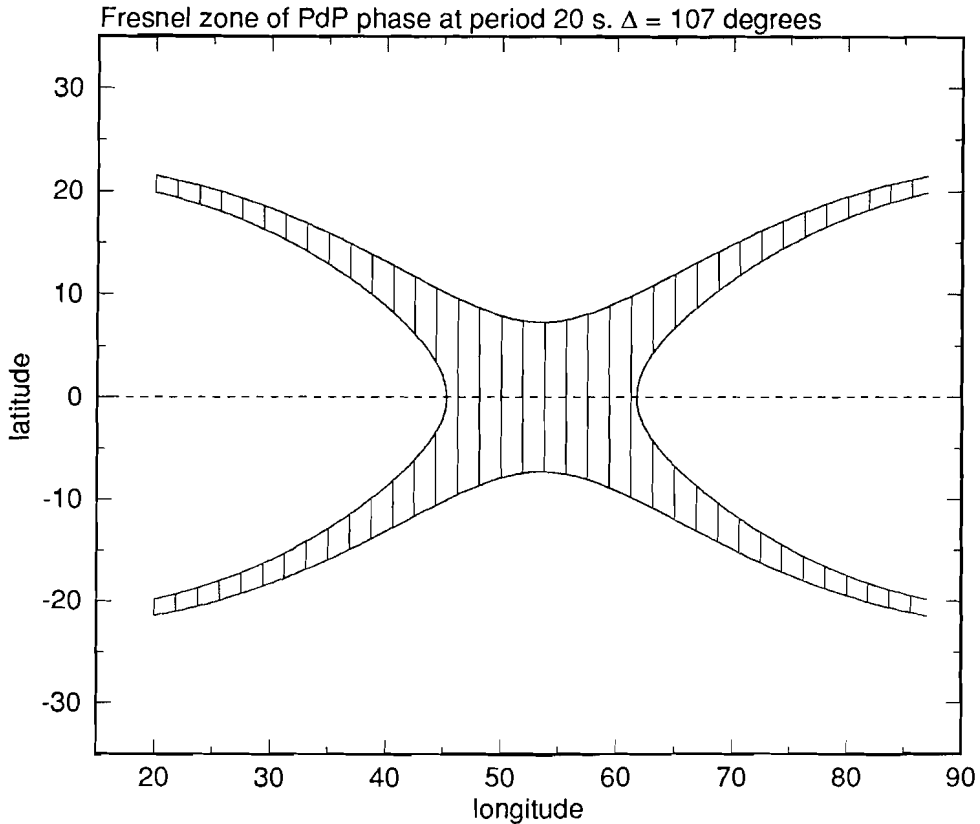


Figure 4. The Fresnel zone for a  $P_{400}P$  phase with a period of 20 s, observed at a distance of  $107^\circ$ . The four tails are due to the minimax nature of  $PP$  (and  $P_dP$ ). The Fresnel zone measures almost  $20^\circ \times 15^\circ$  at the bounce point.

in velocity. The seismic wavelengths at 400 km depth are of the order of 200 km (for a period of 20 s), which is much more than the thickness of the phase change estimated to be 10-20 km (Katsura & Ito 1989). The thickness of the discontinuity is less than a quarter of a seismic wavelength and the data can not distinguish between a gradual increase and a step change in velocity (Richards 1972; Ingate, Ha & Muirhead 1986). Lateral variations in the thickness of the phase change are not seen by the data, unless the thickness exceeds 50 km, which seems unrealistic.

The horizontal resolution of the data can be obtained by calculating the Fresnel zone of the  $P_{400}P$  phase at the reflection point. The frequency content of the data has its maximum around 20 s, quickly falling to zero above 0.1 Hz. The Fresnel zone for a  $P_dP$  phase with a period of 20 s and epicentral distance of  $107^\circ$  is plotted in figure 4. The Fresnel zone is defined here to include all paths that have a time difference with the geometrical path of no more than a quarter period, i.e. in this case between -5 and 5 s. The four long tails of the Fresnel zone reflect the mini-max nature of the  $PP$  phase: the travel time of the ray geometrical path is maximum along the great circle and minimum along a line perpendicular to it.

It appears that  $PP$  and  $P_dP$  phases have an exceptionally large Fresnel zone at the reflection point. A single  $P_{400}P$  phase thus samples a large area on the discontinuity. Structure on the discontinuity that is as far as  $30^\circ$  away from the midpoint between source and station can be sampled by one of the tails and can thus have influence on the amplitude of the  $P_dP$  phase. Assuming that a phase is insensitive to structure with wavelengths either much shorter or much longer than the size of its Fresnel zone, the present data set should yield information on structure on the discontinuity with wavelengths of the order of 1000 km.

### 3. Data processing

Average (absolute) amplitudes of  $P_{200}P$  and  $P_{400}P$  are determined in the time windows containing these phases. To remove the effect of varying source wavelets which to some degree overlap between phases, the traces are deconvolved with the  $PP$  wavelet using the deconvolution method of Langston (1979). The time windows are 35 s long and centered on the arrival times of the phases, as predicted by the PEMc model (see figure 1). As noted above, the  $P_{670}P$  phases are not included in the analysis, as they interfere with  $Pp_{670}p$  (figure 1). The average (absolute) amplitudes of the  $PP$ -precursors are divided by the average (absolute)  $PP$  amplitude. In this way a measure of  $P_dP$  amplitude, relative to  $PP$  is obtained, for both data and synthetics. Figure 5 shows the amplitude ratio  $P_{400}P$ - $PP$  versus epicentral distance, comparing data to synthetics. The observed  $P_{400}P$  amplitudes are distributed around the curve describing the synthetic amplitudes. At distances shorter than  $100^\circ$  the observed amplitude ratio is slightly higher and shows more scatter. This is due to interference with  $Pp_{670}p$  (see figure 1). Although in many records no distinct  $P_dP$  phases are visible, the magnitude of the observed amplitudes is consistent with WKBJ synthetics from PEMc. In the following, only data for distances greater than  $99^\circ$  are used, as for shorter distances information pertaining to the upper mantle below the source and the station creeps into the data. This limits the number of event-station combinations to 29. To obtain a measure of the amplitude perturbation, the observed amplitude ratios are divided by the synthetic ratios. The amplitude perturbations thus obtained contain the a priori assumption that the velocity jump across the 400-km discontinuity is the same as that of PEMc. Other reference models would give different amplitude perturbations. However, the range of perturbations (the ratio of observed maximum and minimum values) is independent of the reference model used. Therefore, in the next sections the range of amplitude perturbations will be used for making inferences on the 400 km discontinuity, rather than the actual perturbations. This has the disadvantage that no absolute measure of the velocity jump at 400 km can be found. The division of the data by synthetic  $P_dP$  amplitudes then only serves to remove the (weak) regional trend, which is present in figure 5. Figure 6 shows the observed amplitude perturbations in the  $P_{400}P$  window as a function of position of the bounce point. The size of the symbol at each reflection point indicates the amplitude of  $P_{400}P$ , relative to the amplitude predicted by PEMc. The symbol in the upper right corner gives the size for no perturbation, i.e. agreement of the observed amplitude with the amplitude predicted by PEMc. The observed amplitude perturbations are plotted at the position of the bounce point, but it must be kept in mind that the amplitude of the  $P_{400}P$  phase is affected by a large region on the discontinuity. Figure 6 shows that the variability in  $P_dP$  amplitude is large and that these variations take place over a short distance; a factor of about 4 over a few degrees distance occurs in a few locations. Another feature of this



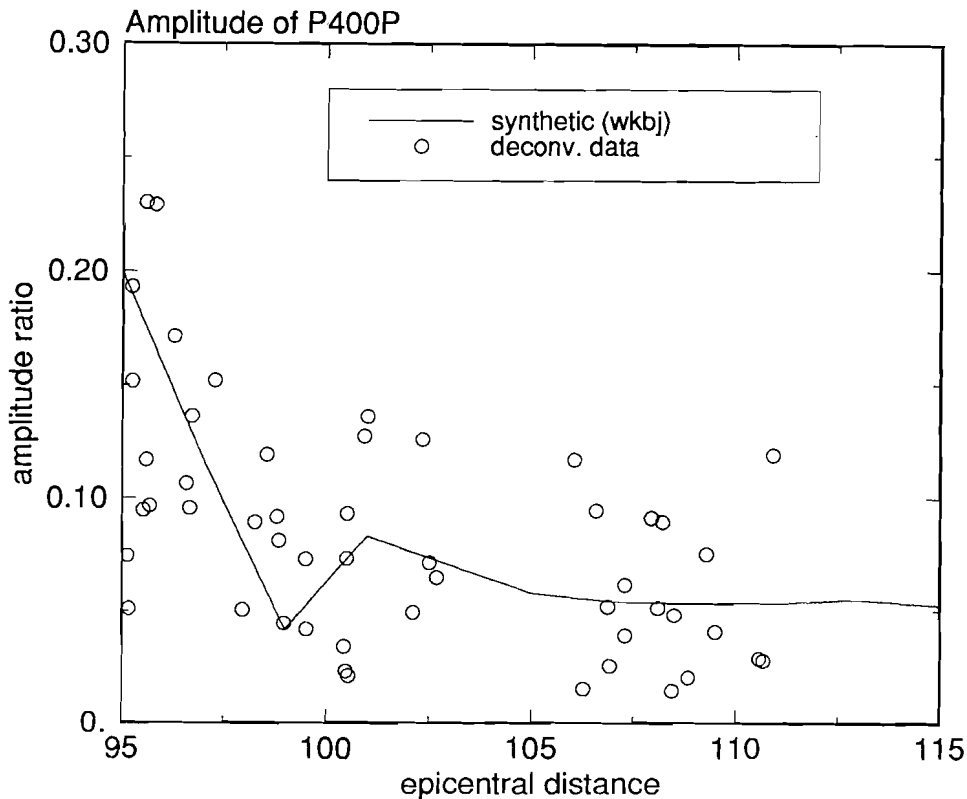
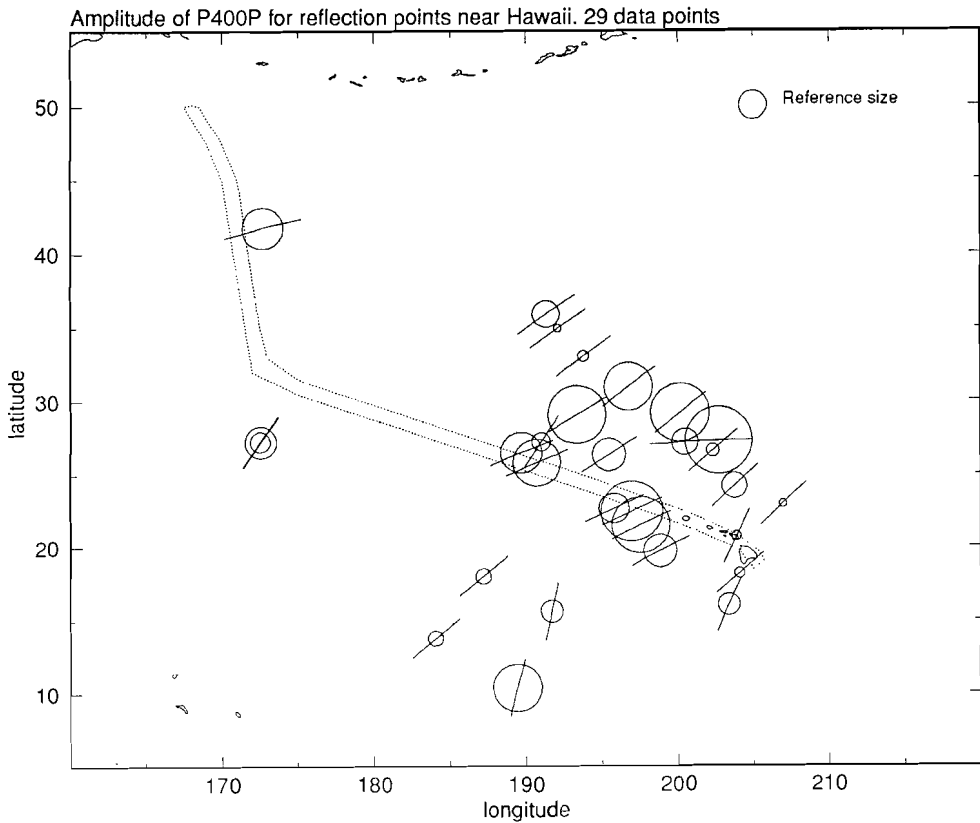


Figure 5. Amplitude ratio  $P_{400}P/PP$  vs. epicentral distance. The observed ratios are distributed around the synthetic values, which include the increase below  $99^\circ$ , which is due to interference with  $Pp_{670}P$ .

figure is that there seems to be no connection between the  $P_{400}P$  amplitudes and the Hawaii hotspot or the hotspot island chain. Figure 6 not only reveals a great variability in the  $P_{400}P$  amplitudes, but also a strong dependence on great circle azimuth, e.g. near  $200^\circ\text{W}$ ,  $30^\circ\text{N}$ , as illustrated in figure 2.

The results for the  $P_{200}P$  time interval show the same variability as those for  $P_{400}P$ , although the variations are smaller. The interpretation of the energy in this interval as entirely due to the  $P_{200}P$  phase is possibly not warranted. Many upper mantle models have several reflectors in the uppermost 200 km of the mantle (see Nolet & Wortel 1989 for a compilation of global and regional models), and underside reflections from such reflectors would overlap and arrive within a short time window.  $P_dP$  phases for  $d$  near 200 km did not show up in the stacks of Shearer (1990) and of Revenaugh & Jordan (1989). The latter authors did find a large gradient at about 100 km depth under the western Pacific, which was interpreted as the onset of partial melting. Bock (1991) confirmed this for  $S$ -to- $P$  conversions beneath Hawaii. Model GCA (Walck 1984), for the upper mantle beneath a spreading centre, has little structure above 200 km depth, but this is the least resolved part



*Figure 6.* Amplitude perturbations relative to PEMc for  $P_{400}P$  as a function of the position of the bounce point. The short line through each point indicates the azimuth of the great circle. The dotted curve is a rough indication of the position of the Hawaii-Emperor hot spot chain. The symbol at the upper right gives the reference size: a symbol of this size indicates that the observed  $P_{400}P$  amplitude agrees with the amplitude predicted by PEMc. Note that the amplitude may vary strongly over short horizontal distances; the azimuth dependence is apparent for closely spaced bounce points.

of that model. It is possible that several bottom-side reflections arrive within the  $P_{200}P$  interval; travel time differences among such phases may be completely responsible for the amplitude variations. The results for the  $P_{200}P$  interval are therefore subject to much greater uncertainty than those for the  $P_{400}P$  phase and these data will not be used in the following sections.

No results can be obtained for the underside reflection of the Moho, as this phase arrives too close to  $PP$  to be resolved. It is expected that the Moho depth does not vary much in oceanic areas, except in the vicinity of the Hawaii hot spot swell. Refraction studies near Hawaii showed that the Moho is elevated parallel to the sea floor (Shor & Pollard 1964). Both the wavelength and amplitude of this elevation is probably too small to have an appreciable effect on the Moho reflection and it is assumed that there is no effect of Moho undulations in the data.

In the following sections only the results for the 400 km discontinuity are studied and the variations in amplitude of  $P_{400}P$  phases are explained as due to topography on the 400 km discontinuity, i.e. the amplitude perturbations are explained as focusing and defocusing. Combining the observations of figure 6 with the two characteristics that were derived from the data in the previous section, there are now four important characteristics that acceptable models for the 400 km discontinuity have to be able to explain:

- (i) The range of focusing. The maximum and minimum observed focusing values differ by about a factor of 8.
- (ii) The azimuth dependence of the  $P_{400}P$  amplitude. Figures 2 and 6 show that this dependence is strong.
- (iii) The variability of the focusing. Figure 6 shows that a factor of 4 for bounce points only a few degrees apart is observed.
- (iv) The character of the seismogram between  $P$  and  $PP$ . Apart from explaining focusing or defocusing of  $P_{400}P$ , the model must also give an explanation for the reverberatory behaviour of this part of the seismic signal.

The data in figures 2 and 3 indicate that the resolving power of the method will not be high, as no distinct arrivals can be seen and the  $P_dP$  phases are weak. The data can be used, however, to find constraints on wavelength and amplitude of topography on the 400-km discontinuity, as is shown in the next sections.

#### 4. Possible sources of error

A number of sources of error exist that may bias the observed focusing factors or produce a systematic error.

*Radiation pattern.* As the  $PP$  and  $P_dP$  waves leave the source at a different take-off angle, the radiation pattern at the source may be important. However, the difference in take-off angle between  $P_{400}P$  and  $PP$  for an epicentral distance of about  $100^\circ$  is less than  $0.5^\circ$ . Only close to a node in the radiation pattern will this cause a significant error in the  $P_{400}P$  -  $PP$  amplitude ratio. In such a case,  $PP$  itself will be small. The data have been checked and do not contain anomalously small  $PP$  waves.

*Low Q in the LVZ.* The  $PP$  wave path traverses the low-velocity zone (LVZ) four times, whereas the  $P_{400}P$  phase only passes the LVZ twice. The LVZ is generally considered as a zone of large damping, and as damping is not considered in the calculation of the synthetic wave forms, this could be a source of systematic error. However, due to the long wavelengths of the data (at 100-200 km depth typically about 150-200 km), the effect of low Q in the LVZ is small: a 100 km thick LVZ with Q as low as 100 would introduce a relative error of only 5% in the focusing values.

*Station dependence.* A possible site-related bias can arise from long reverberations under the station. The reverberations produce long coda following each phase, causing overlap of wave trains. Reverberation may depend on angle of incidence, so  $PP$  may excite a reverberation different from that of  $P$ . To reduce this source of error, all data that show a long, reverberatory  $P$  or  $PP$  coda are excluded from the data set.

*Focusing due to other structure.* Lateral heterogeneity other than displacement of

discontinuities may cause focusing or defocusing of  $PP$  and  $P_{400}P$ . However, the ray paths of  $PP$  and  $P_{400}P$  differ only beneath the reflection point, above the reflection depth of  $P_{400}P$ . Only in this region can heterogeneity affect  $PP$ , and not  $P_{400}P$ . Displacement of interfaces has negligible effect on  $PP$ , as the travel time perturbations thus induced are small. Topography at the free surface may have a noticeable effect. However, a surface displacement of 3 km at the bounce point induces a travel time perturbation of roughly 1 s, which is too small for the periods involved here (longer than 15-20 s) to cause significant focusing. Volume heterogeneities beneath the bounce point, finally, are assumed to have a small effect on the amplitude of  $PP$ . This is illustrated by the records in figure 2, in which the  $P_dP$  phases for the same bounce point have different character.

*Overlap with other  $P_dP$  phases.* Figure 5 shows a marked increase in the observed relative amplitude below  $99^\circ$ . This is due to overlap with the  $Pp_{670}p$  phase, as is clear from the WKBJ synthetics. These data have not been used. It is possible that there is also overlap of  $P_{400}P$  with the earlier arriving  $P_{670}P$  phase. The difference in arrival time between the phases is about 40 s in PEMc, which is about as long as the average  $P$  wavelet. Several authors have reported the existence of a reflector near 500 km depth (see Shearer 1990, and references therein; Stammer et al. 1991). As such reports are quite uncommon and phases related with this discontinuity appear only faintly, it is assumed that this reflector, if present, has minor effect on the data. As  $P_dP$  are minimax phases, related phases may arrive early. For  $P_{400}P$ , this implies that  $P_{200}P$  and  $P_{670}P$ -related arrivals may interfere. This problem will be addressed in the sections below.

## 5. Focusing of $P_dP$ phases

Focusing of energy at the bounce point of  $P_dP$  phases may account for the anomalous observations presented in the previous sections. Whitcomb (1973) suggested this mechanism to explain the intermittent observation of such phases. Davis et al. (1989) computed the focusing of precursors to  $P'P'$  using Gaussian beams. From their 2-D experiments it appeared that topography on the 670-km discontinuity with a wavelength of 1000 km and an amplitude of 10 km can affect the amplitude of  $P'_{670}P'$  by a factor of 2. Revenaugh & Jordan (1989) invoked topography of upper mantle discontinuities to explain the variation in arrival times of  $ScS$ -related reverberations in the upper mantle. In this section a procedure for estimating the focusing of  $PP$  precursors by full 3-D topography on the upper mantle discontinuities is presented.

An adequate way of studying the effect of topography on a discontinuity on the amplitude of a  $P_dP$  phase is the Kirchhoff-Helmholtz integral, which is essentially a sum over all possible paths from the source to some surface (the discontinuity) and from this surface to the receiving station (see e.g., Frazer 1987; Sen & Frazer 1987). Symbolically,

$$A_d(\omega) = i\omega A_P(\omega) \int_{\Sigma} R(\omega, \mathbf{r}_s, \mathbf{r}) P(\mathbf{r}_s, \mathbf{r}) P'(\mathbf{r}, \mathbf{r}_r) e^{i\omega T(\mathbf{r})} dS \quad (1)$$

where  $A_d$  is the spectral amplitude of the  $P_dP$  wave;  $\Sigma$  is the integration surface;  $R(\omega, \mathbf{r}_s, \mathbf{r})$  is the generalized (frequency and slowness dependent) reflection coefficient;  $A_P$  contains the amplitude spectrum of the  $P$  wave;  $P$  and  $P'$  describe the change in amplitude for the propagation from the source at  $\mathbf{r}_s$  to the integration point at  $\mathbf{r}$  and from  $\mathbf{r}$  to the station at  $\mathbf{r}_r$ , respectively; the exponent describes the phase shift due to travel time  $T$  along the

propagation path. The factor  $\omega$  arises from the fact that the integral behaves as  $1/\omega$ ; this can be seen by applying a stationary phase approximation (see Bleistein 1984). The physical meaning of this is that the integration works as a low-pass filter; the factor  $\omega$  compensates for this.

A theoretical focusing factor can be calculated that contains the dominant contribution from displacements of the reflecting interface. This focusing factor is the spectral amplitude for a perturbed interface over the amplitude for a flat (unperturbed) reflector

$$F(\omega) = \frac{|A_d^{perturbed}(\omega)|}{|A_d^{flat}(\omega)|} \quad (2)$$

This factor is frequency dependent. The advantage of the calculation in the frequency domain is that frequency-dependent effects are taken into account. Such effects are expected for a reflection at a reflector with laterally varying properties. A stationary phase approximation would not be appropriate, as that performs a high frequency approximation and removes frequency-dependent effects.

In principle, the integration surface  $\Sigma$  in (1) extends to infinity, or, as in this case for spherical geometry, includes the entire sphere. However, in this study the region of interest is the neighbourhood of the bounce point and the integration is limited to this region. Near the bounce point, the integrand is stationary. Away from the stationary point, the integrand becomes oscillatory. (In the stationary phase approximation, it is assumed that the contribution from this oscillatory domain vanishes.) The integration limits are chosen in such a way that the integrand is well into the oscillatory domain. In the calculations of the next section, the integration area used was an area of  $70^\circ \times 70^\circ$ , centered on the midpoint between source and station. A taper function is applied to the edges the integration area, to avoid a dependence of the integration on the choice of integration limits and to suppress cut-off artifacts.

Equation (2) gives the focusing at each frequency  $\omega$ . This is not necessarily the same as the  $P_dP$  amplitude increase or decrease observed in the seismogram, as  $F(\omega)$  is frequency dependent and this dependence may be quite strong. For multiple stationary points in the area of integration (and thus multiple  $P_dP$  arrivals) equation (1) contains a sum over the amplitude of these arrivals. If all stationary points are within the Fresnel zone, the arrivals add constructively and the focusing value is a measure of amplitude in the seismogram. However, some examples in the following sections show that significant amount of energy may arrive well ahead and after the theoretical  $P_dP$  phase. In such cases, the value predicted with equation (2) may not reflect the observed  $P_dP$  amplitude. To check this, equation (1) is used to calculate  $P_dP$  wave forms: for each frequency both the unperturbed and perturbed spectra  $A_d(\omega)$  are computed and multiplied with the desired  $PP$  wave spectrum; an inverse Fourier transform then gives the unperturbed and perturbed  $P_dP$  waveforms.

## 6. Models of topography

A number of models of topography was used to try to explain the observed characteristics of the data. It was attempted to find ranges of scale lengths and amplitudes that are permitted by the data.

*Single sinusoidal deflection.* If it is assumed that the origin of the Hawaii hotspot lies deep within the Earth, perhaps on the core-mantle boundary as several studies suggest

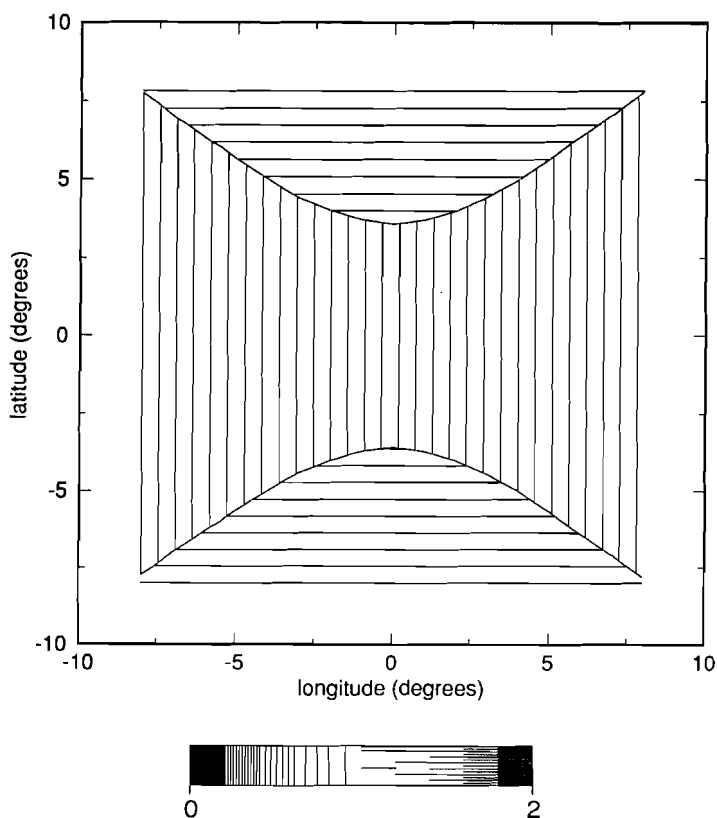


Figure 7. The focusing of a 20 s  $P_{400}P$  phase predicted by a simple model of the deflection of the 400-km discontinuity caused by a plume. The single sinusoidal deflection has a width of 500 km, and is 20 km deep in the centre. The figure shows the focusing as a function of the position of the bounce point, relative to the 'hot spot', which is located at the origin. The epicentral distance is  $100^\circ$ .

(Yuen & Peltier 1980; Richards, Hager & Sleep 1988; Liu & Chase 1991), then the plume affects the upper mantle discontinuities. Richards et al. (1988) concluded from geoid data that the channel feeding hot mantle material to the hotspot should be typically a few hundred kilometers wide in the lower mantle, with material in the pipe  $200^\circ\text{--}700^\circ\text{K}$  hotter than the ambient mantle.

With the assumption that such a plume would rise into the upper mantle, the effect of such a continuous plume rising through the 400-km discontinuity can be roughly calculated. Using a Clapeyron slope  $\gamma = 3.6\text{ MPa/K}$  (Liu et al. 1991), the displacement of the discontinuity is about 10 km for every 100 K temperature difference between the plume and the mantle. The phase change at 400 km is exothermic, so inside the plume the discontinuity is at greater depth. The width of the perturbed zone follows from a simple conduction calculation. If it is assumed that the plume has been in place for the last 70–80 Ma (which is the age of the oldest volcanic islands of the Emperor chain), and using  $\kappa = 0.1\text{ s/mm}^2$  for the thermal conductivity, a value of 500 km is found for the diameter of the thermally

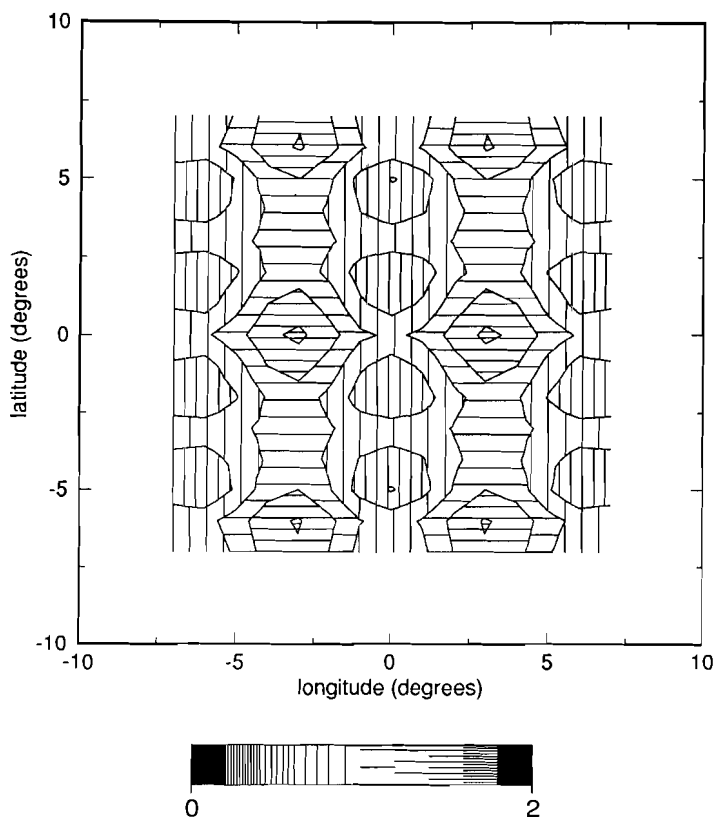


Figure 8. Same as figure 7, now for a harmonic topography. The wavelengths of the topography are  $1000 \times 1000$  km, the (peak-to-peak) amplitude 40 km. The great circle is at an angle of  $45^\circ$  with the symmetry axes of the topography.

affected zone around the plume. This displacement is approximated with a single sinusoidal minimum with an amplitude of 20 km and a width of 500 km. The results were obtained by moving the bounce point relative to the topography, keeping the great circle azimuth constant. The epicentral distance was  $110^\circ$  (this is the case for figures 7-16) and the results are valid for a period of 20 s. The 'hot spot' in figure 7 is located at the origin. The results in this figure can be compared directly with figure 6. This figure proves that this simple model does not explain any of the characteristics of the observations: both the amplitude and the variation in amplitudes are small (6% variation). As the model is symmetric, there is no dependence on great circle azimuth. Assuming a hotter plume, with a temperature difference of 700 K (suggested by Richards et al. (1988)) does not improve the results. A hotter plume has a wider thermally affected area surrounding it, apart from an increased displacement of the phase transition. The net effect to the results in figure 7 would be that although the (de)focusing may increase, the variations as a function of position relative to the great circle become slower.

*Series of sinusoids.* Recent results in numerical convection experiments, in which the

upper mantle phase changes have been included, show that the notion of a continuous plume throughout the mantle may not be realistic. The effect of the phase change at 670 km is that material rising up through it becomes more buoyant and tends to rise more rapidly; numerical experiments show that this has the effect of periodic decapitation of the plume (Liu et al. 1991) and that in the upper mantle no continuous plume is present, but rather a series of blobs of hot material. Recent work of Machetel & Weber (1991) also shows this behaviour which leads to intermittent exchange of material between upper and lower mantle.

If such a type of hot spot feeding mechanism were active at Hawaii, then the 400-km discontinuity might be modeled with a series of bumps under the chain; these would be the thermal remnants of the blobs of hot material that once passed. For simplicity, the individual sinusoids are taken identical to the first model, i.e. 500 km wide and 20 km deep. The results of this model (not shown) indicate that this model also fails to explain the data characteristics, although the amplitudes now vary over about 30%. This model does not have rotational symmetry, so the focusing depends on the great circle azimuth. This dependence is, however, too small to explain the data.

*Harmonic topography.* If a single deflection or a short series of deflections can not explain the data, perhaps a large number of sinusoids can. Figure 8 shows the focusing of a model of harmonic topography with a wavelength of 1000 km in both longitude and latitude directions and an amplitude of 20 km; the results are for a  $P_dP$  phase with a period of 20 s. To find the maximum focusing possible with such a model, the great circle intersects the topography at an angle of  $45^\circ$  with the symmetry axes. The structure that best focuses a  $P_dP$  phase is parallel to the surface of equal travel time, so that all points radiate in phase. For a  $P_dP$  phase, this surface is a saddle structure, which is reflected in the shape of the Fresnel zone (figure 4). This optimum focusing configuration is achieved locally for a harmonic topography with symmetry axes at an angle of  $45^\circ$  with the great circle. Each maximum in figure 8 corresponds with such a configuration. The periodicity in the topography is reflected in the periodic behaviour of the focusing, which has the same wavelength as the topography. Although the difference between maximum and minimum focusing is smaller than the observed difference, this model does explain the rapid variations in the data. Over a distance of less than 5 degrees, the focusing changes from 0.6 to 1.4. Figure 9 shows the dependence on great circle azimuth for the point at (lon,lat)=( $0^\circ$ ,  $0^\circ$ ) in figure 8. It shows that this harmonic model also explains the azimuth dependence of the focusing: over a short range of azimuth (only  $10^\circ$ ), the focusing varies by a factor of 4. The jaggedness of the curve in this figure can be explained by stationary points near the edges of the Fresnel zone: changing the azimuth is equivalent to rotating the Fresnel zone relative to the model of topography. This has the effect of intermittently turning stationary points near the edges of the Fresnel zone 'off' or 'on', as they fall outside or inside the Fresnel zone (remember the long tails in figure 4). Note that this is not a cut-off effect: the area of integration in equation (2) extends well beyond the first Fresnel zone.

As this model generates many secondary stationary points, it is interesting to consider the waveforms produced by this model. Figure 10a shows the perturbed  $P_dP$  phase due to the harmonic model described above, for a bounce point at (lon,lat)=( $-3^\circ$ ,  $0^\circ$ ). Whereas the results in figure 8 give a focusing by a factor of 1.4 (the focusing is about 0.5), figure 10a shows a perturbed  $P_dP$  phase (dashed line) with a maximum amplitude comparable to the unperturbed phase (solid line). The reason for this discrepancy is shown in figure 10b,



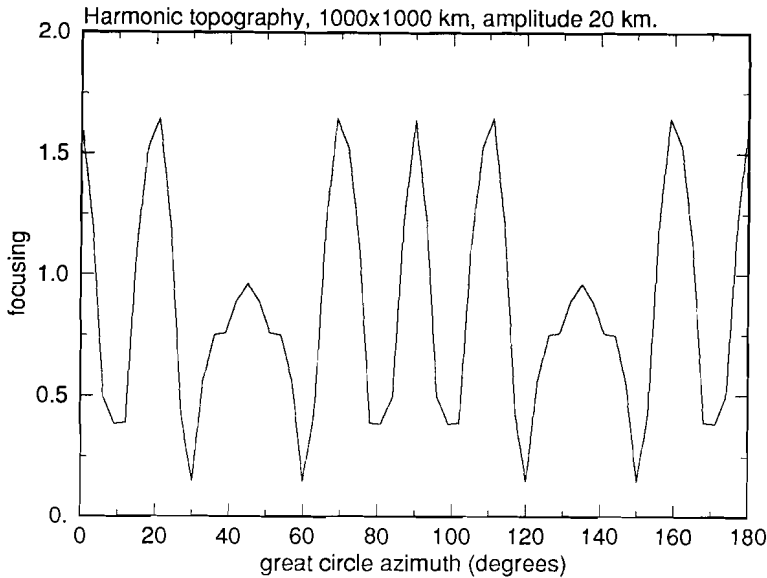


Figure 9. Dependence of the focusing on the azimuth of the great circle for the model used in figure 8. Strong variations are predicted for small changes in azimuth. See text for an explanation of the jaggedness of the curve.

where the focusing factor of the perturbed  $P_dP$  phase (dotted curve) and the unperturbed  $P_dP$  phase (solid curve) are plotted. The focusing factor for the perturbed phase varies strongly within the bandwidth of the  $P_dP$  phase. The multiplication of the focusing factor with the  $PP$  spectrum and the FFT to the time domain averages out the variations in (de)focusing as a function of frequency. This leads to a main pulse which is almost unperturbed in amplitude; the effect of the secondary stationary points can be clearly seen in the significant amount of energy arriving before and after the main pulse. The perturbed wave form in figure 10a does explain the fourth characteristic of the data. The  $P_{400}P$  wavelet is smeared out by reflection at the undulating discontinuity and this explains why often no individual  $P_{400}P$  can be identified in the records. Assuming that such an effect is present for other discontinuities as well, overlap between different  $P_dP$  phases generates the reverberatory  $P$  coda, in which pulse-like arrivals cannot be seen.

The  $P_dP$  wave form for a harmonic topography with wavelengths  $2000 \times 2000$  km and amplitude 20 km is shown in figure 11a. This topography does not create as many stationary points outside the Fresnel zone, as is apparent from the energy before and after the main pulse. The pulse has longer periods than the original wavelet; this is caused by the lower frequencies being less defocused than higher frequencies (figure 11b). As for the previous case, the focusing predicted for a frequency of 0.05 Hz (20 s) is not a reasonable indication of the actual amplitude of the  $P_dP$  phase. This topography (wavelength  $2000 \times 2000$  km) does not explain the first two data characteristics: range and spatial variability of focusing.

*Random topography.* Although the harmonic topography explains most of the data characteristics, it is interesting to find the focusing predicted by a model of random topography on the 400-km discontinuity. A quasi-random model was constructed by averaging a grid

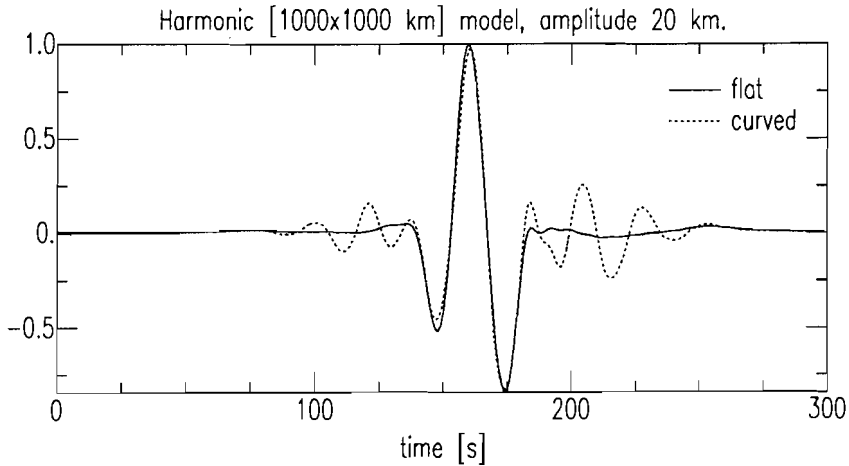


Figure 10a.  $P_{400}P$  wave form calculated for the model of figure 8 (dotted line) and unperturbed  $P_{400}P$  wave (solid line). The topography has strongly affected the  $P_{400}P$  wave form, although the maximum amplitude is almost unchanged. Strong leading and trailing oscillations appear.

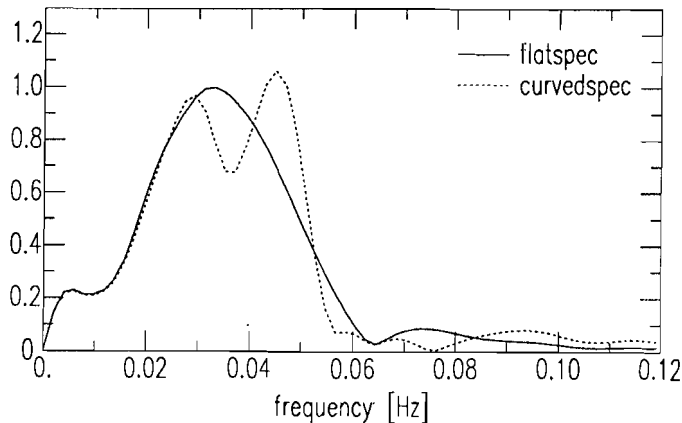


Figure 10b. Spectra of the wave forms of figure 10a. Solid line is the spectrum for the unperturbed wave, the dotted line for the perturbed wave. Note that the focusing factor for the perturbed  $P_{400}P$  wave varies strongly within the bandwidth of  $PP$ .

of random numbers over a moving window of specified size. The resulting smoothed random model was scaled to a specified r.m.s. value. (Such a random model is described by a sinc-function behaviour in the  $k_x$ - $k_y$  domain.) The focusing of a 20 s  $P_{400}P$  phase predicted by a model with averaging window of 1000x1000 km and an r.m.s. amplitude of 15 km are presented in figure 12. This model can also explain the data characteristics, including the dependence on great circle azimuth (figure 13). The focusing for this particular random model ranges from 0.15 to 1.65, which compares quite well with observed amplitude

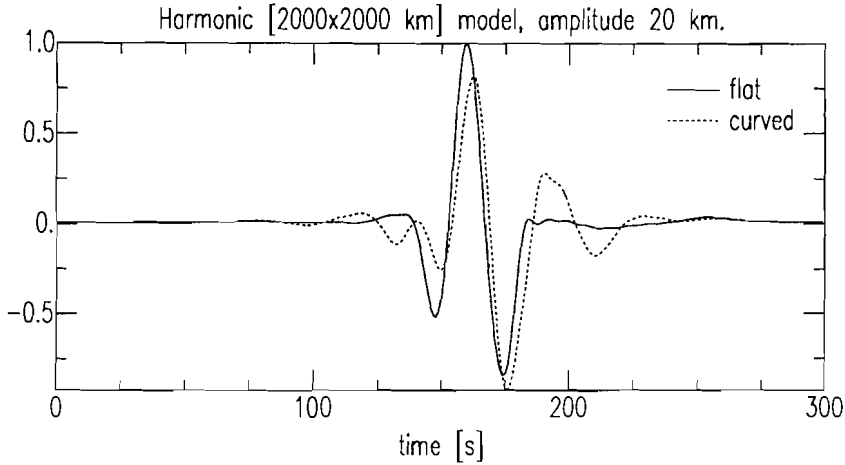


Figure 11a. As figure 10a, now for a harmonic model with wavelengths of 2000x2000 km and amplitude 20 km.

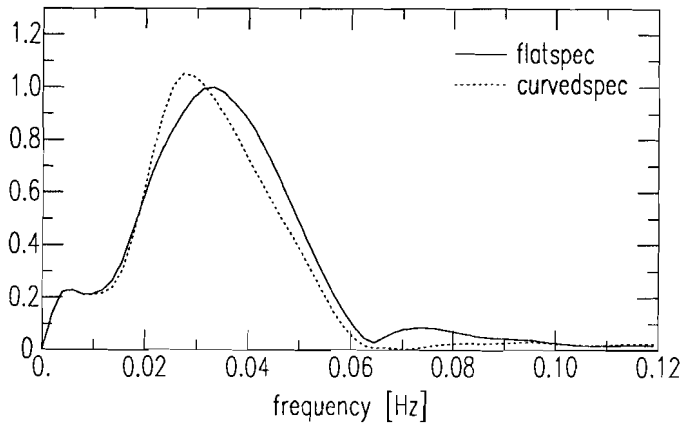


Figure 11b. The spectra of the wave forms in figure 11a.

range, and variations take place over short distances. A random model with averaging lengths 500x500 km and r.m.s. value of 15 km predicted the same amplitude variations; these variations take place over shorter distances, a change in focusing of a factor of 4 over 3° is possible. Random models with a r.m.s. amplitude smaller than about 15 km or averaging lengths longer than about 1500 km do not predict amplitude variations comparable to observed variations.

An example of a  $P_dP$  wave form produced by the random model used for figure 12 is shown in figure 14. The bounce point is at -6° longitude, -6° latitude: near the maximum of figure 12. Although the focusing depends on frequency, this dependence is not so strong that the perturbed phase has the same amplitude as the unperturbed phase. Figure 10b

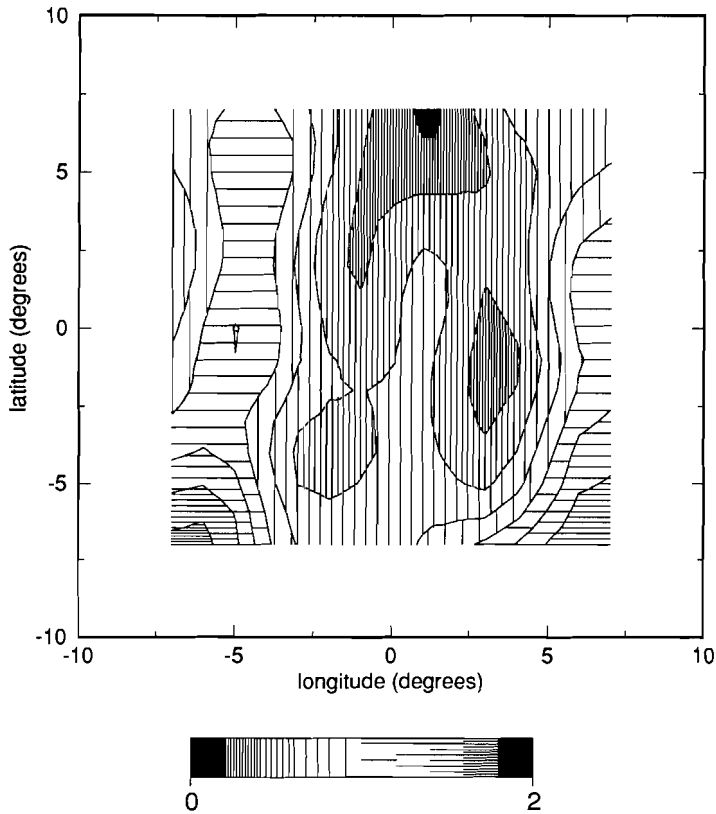


Figure 12. As figure 7, now for a random model, described by an averaging window of  $1000 \times 1000$  km and r.m.s. amplitude 15 km. The focusing ranges from 0.15 to 1.65.

shows that the frequency dependence of the focusing factor for the harmonic model is so strong that after convolution with the  $PP$  spectrum the maximum amplitude of the  $P_{400}P$  wave form is almost the same as that of the unperturbed phase. The defocusing predicted for one frequency is not reflected in the wave form. The frequency dependence of the focusing factor is much weaker for the random model, as can be seen in figure 14: the focusing predicted for  $T=20$ s is reflected in the perturbed wave form. Notice that secondary stationary points cause a tail following the main arrival and to a lesser extent a precursor before the main arrival. Figure 15 shows a  $P_{400}P$  phase from a random model with averaging lengths of  $500 \times 500$  km. In this case, the  $P_{400}P$  phase is seriously distorted, large-amplitude oscillations appear, arriving later than the unperturbed phase (solid curve). These earlier and later arrivals explain the reverberatory characteristic of the data. The overlap of perturbed  $P_dP$  phases in the seismogram causes the reverberatory behaviour and makes the identification of individual arrivals difficult. This overlap of  $P_dP$  phases introduces an error into the calculation of the amplitude ratio  $P_dP-PP$ . It is difficult to quantify this error; the models used here did not generate significant energy well ahead or after the main  $P_dP$  pulse.

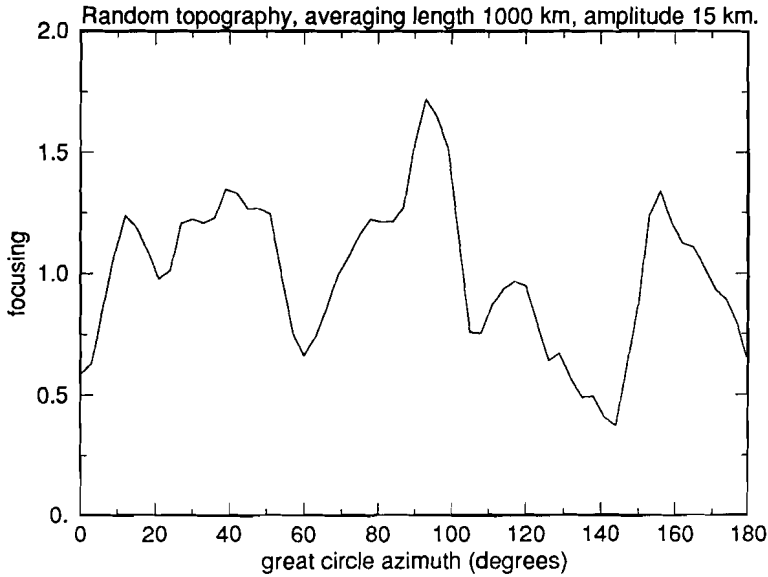


Figure 13. Azimuth dependence of focusing factor for a bounce point at  $(0^\circ, 0^\circ)$  in the model used in figure 12. Strong variations are found for small changes in azimuth.

*Variable reflection coefficient.* Another cause for the variability of the  $P_dP$  phases can be that the reflection coefficient of the discontinuity varies laterally. This can be caused by a varying jump in velocity. It is interesting to find its effect on  $P_dP$  amplitudes as a check of the uniqueness of the topographic explanation of the  $P_dP$  amplitudes. A laterally varying reflection coefficient was modeled by a harmonic function  $R = R_o(1 + A \sin \kappa \theta \sin \lambda \phi)$ , with wavelengths  $\kappa$  and  $\lambda$  of 1000 km. With  $A = 0.7$ , the reflection coefficient varies between  $0.3R_o$  and  $1.7R_o$ . With this strongly oscillating reflection coefficient,  $P_dP$  amplitude variations from 0.7 to 1.4 are found. This is too small to explain the data. Therefore, it is unlikely that laterally varying velocity contrast is the cause of the observed  $P_dP$  amplitude variations.

*Best models.* The class of models that best explains the character of the observed  $P_{400}P$  phases is the (quasi-)random model. All four characteristics of the data can be explained. An estimate of the possible range of wavelength and amplitude of the topography can be obtained by calculating the focusing predicted by a number of models, each with different averaging length and r.m.s. topography. For each model, the focusing as a function of the position of the bounce point is calculated, for periods of 20 s, 25 s and 33 s, spanning the interval of dominant periods in the data. A model is considered acceptable if, for each period, it predicts a factor of about 4 between the extremes of (de)focusing and if such variations take place within a few degrees. It must be noted that although the models are random, they represent only one realization. No averaging was done over a large number of random models.

The range of acceptable models is presented in figure 16. The combinations of r.m.s.

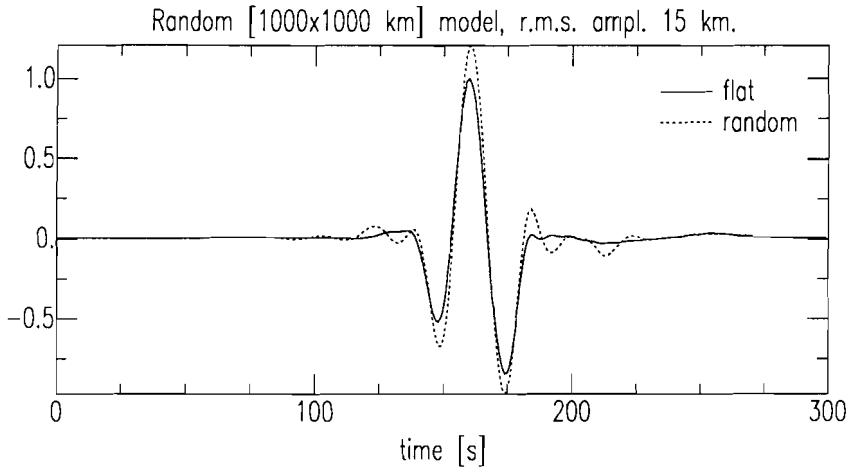


Figure 14.  $P_{400}P$  wave form predicted by the model used in figure 12, for a bounce point at  $(\text{lon, lat}) = (-6^\circ, -6^\circ)$ , near the maximum in figure 12 (dotted line). Again leading and trailing oscillations appear.

amplitude and averaging length that satisfy the requirements fall in the grey area. The borders of the parameter domain that defines acceptable models is not sharply defined. No model with an r.m.s. amplitude of 10 km or smaller could explain the data; r.m.s. amplitudes of 25 km and 30 km were strongly dominated by defocusing and thus largely failed to explain focusing.

The contoured area illustrates the relation between wavelength and amplitude: topography of longer wavelengths must have higher amplitude in order to accomplish the same focusing as short wavelengths. The reason for this is that the curvature of the topography is the important parameter: it determines how well the topography parallels the  $P_dP$  travel time surface. This also facilitates the interpretation of figure 16. The space of averaging length-rms amplitude can roughly be divided into four regions. At low amplitudes (indicated by the letter B in figure 16) the models are too smooth to create appreciable focusing. At high amplitudes (region A) the topography becomes so steep that energy is scattered in all directions; these models mainly defocus  $P_dP$  phases. The contoured area is in between these regions. The contour is closed at long averaging lengths and high amplitudes (region D) because such models can not explain rapid  $P_dP$  amplitude variations, although they do explain the observed amount of (de)focusing. At low amplitudes and short averaging lengths (region C) the models generate amplitudes fluctuations on a scale much smaller than the size of the Fresnel zone. Integration over this zone averages out these variations, leaving only small variations. The results suggest that the optimum averaging lengths lie between about 500-1500 km, with the r.m.s. height about 15-20 km. The wavelengths (in contrast to the averaging length used to describe the models) contained in the random models range from about 500 to 4000 km. The wavelengths actually contributing to the focusing can be estimated by inspection of the focusing as a function of bounce point position (figure 12) for all acceptable models. As for the harmonic model (figures 7 and 8), the distance between maxima and minima of focusing reflects the wavelengths of the topography. The

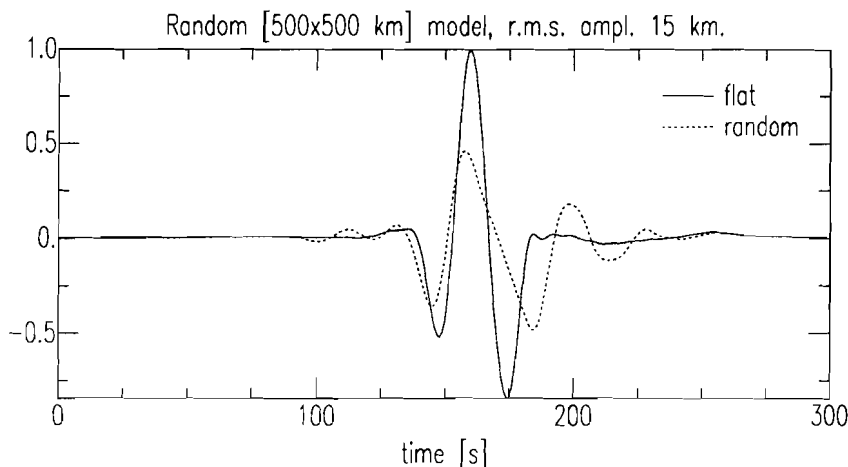


Figure 15.  $P_{400}P$  wave form for a random model with averaging window of 500x500 km and r.m.s. amplitude 15 km (dotted line), and unperturbed wave form (solid line).

range of wavelengths thus appears to be roughly 600-2000 km. This is also the range that can be derived from the variations in the observed  $P_{400}P$  amplitudes in figure 6.

## 7. Conclusions

A method has been presented to find a measure of the topography on upper mantle discontinuities. The result is not a deterministic map of the topography, but a range of combinations of wavelength and amplitude that are permitted by the data. If many data were available, imaging of the topography on the 400 km discontinuity would in principle be possible. From the results presented here a number of conclusions can be drawn.

*Amplitude/wavelength of topography.* The models that best explain the observed amplitude variations of long-period  $P_{400}P$  have rms amplitudes of about 15-20 km. The wavelengths of the topography are about 700-2000 km. These results are obtained with a number of assumptions (e.g., no scattered energy, no interference between phases). This means that the range found for the amplitudes of the topography must be regarded as an upper limit. The range of wavelengths found here is better constrained, if topography on the discontinuity accounts for the lateral variations of  $P_dP$  amplitude.

The range of wavelengths and amplitudes found in the present study compares well with those reported in a number of different studies of varying discontinuity depth in the upper mantle. Revenaugh & Jordan (1989), who studied the 670-km discontinuity with  $ScS$ -related phases, found wavelengths in the range of 300-1500 km and peak-to-peak amplitudes of about 20 km. Barley et al. (1982) and Bock & Ha (1984), using  $P$ -to- $S$  converted phases, reported variations in the depth of the 670-km discontinuity of the order of 50 km, near the Izu-Bonin and Tonga subduction zones. Richards & Wicks (1990) arrived at somewhat less warping of this discontinuity beneath the Tonga trench (20-30 km).

*$P_dP$  wave forms.* The synthetic wave forms in figures 10a, 11a, 14 and 15 suggest that topography on the upper mantle discontinuities has the effect of smearing out the  $P_dP$

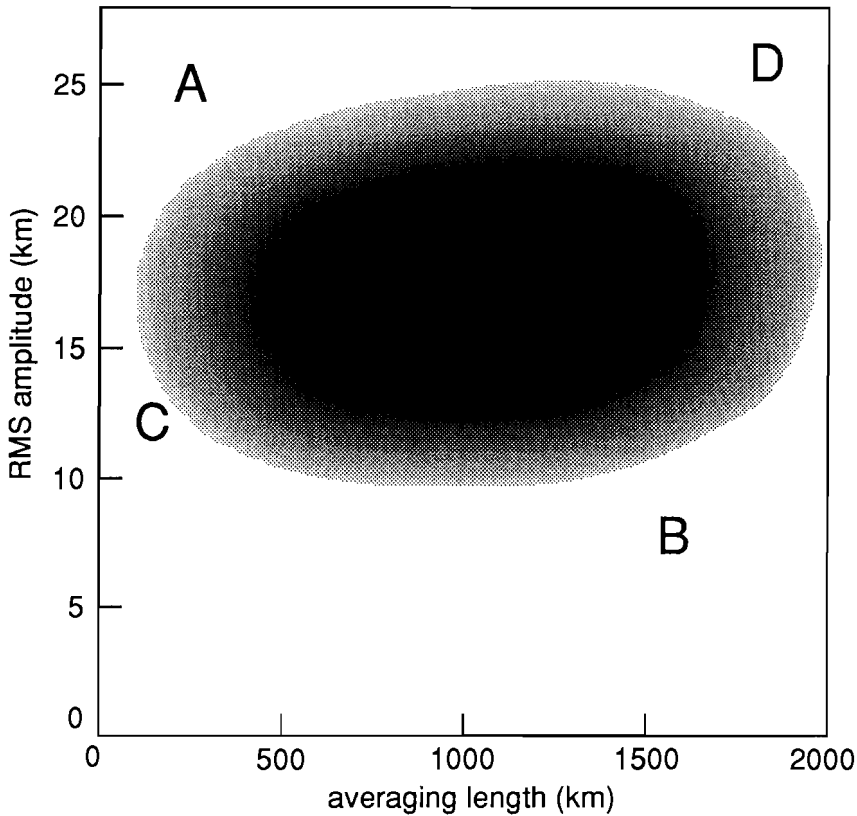


Figure 16. Range of acceptable averaging lengths - r.m.s. amplitude combinations. See text for explanation of the regions A-D.

wave train. In some cases, the earlier or later arriving energy can have amplitudes as large as half of that of the maximum. This results in longer  $P_dP$  wave forms that partly overlap, creating a reverberatory interval between  $P$  and  $PP$ , in which individual arrivals are difficult to observe. This explains the character of the data in figure 3 and also explains the lack of well-developed  $P_dP$  phases reported in the literature (Wajeman, 1988; Davis et al., 1989). This also renders this part of the seismogram unusable for wave form inversion, as the wave forms may be seriously distorted upon reflection at the discontinuity.

Other discontinuity-related phases are reported in the literature: the long-period world-wide stacks of Shearer (1990) contain clear examples, such as  $PdsP$  and  $Pp_d p$ . These phases all have a multiplicity greater than one, and are not obscured by stronger phases.

*Defocusing and focusing.* Figures 8 and 12 show that the different models (harmonic and random) mainly predict defocusing of the bottom-side reflections. Even the harmonic topography does not have an equal distribution of focusing and defocusing. The average focusing in a window as long and wide as the topographic wavelength in figure 8 can be as low as 0.75. This means that averaging (stacking) many  $P_dP$  phases may lead to a reflection coefficient which is biased by topography.



*Plume structure beneath Hawaii.* The data presented here can not determine whether or not a plume structure is present under Hawaii. A simple model of topography induced by a plume shows that the expected focusing is small; such signal will probably drown in stronger signal from larger scale topography.

**Acknowledgment.** Discussions with G. Nolet are greatly appreciated. S. van der Lee is thanked for critically reading the manuscript. The comments of two anonymous reviewers are appreciated.

## References

- Barley, B.J., J.A. Hudson, and A. Douglas, 1982. *S* to *P* scattering at the 650-km discontinuity, *Geophys. J. R. astron. Soc.*, **69**, 159-172.
- Bleistein, N., 1984. *Mathematical methods for wave phenomena*, Academic Press, Orlando.
- Bock, G., 1991. Long-period *S* to *P* converted waves and the onset of partial melting beneath Oahu, Hawaii, *Geophys. Res. Lett.*, **18**, 869-872.
- Bock, G., and J. Ha, 1984. Short-period *S-P* conversion in the mantle at a depth near 700 km, *Geophys. J. R. astr. Soc.*, **77**, 593-615.
- Bolt, B.A., 1970. *P<sub>d</sub>P* and *PKiKP* waves and diffracted *PcP* waves, *Geophys. J. R. astron. Soc.*, **20**, 367-382.
- Davis, J. P., R. Kind, and I. S. Sacks, 1989. Precursors to *P'P'* re-examined using broad-band data, *Geophys. J. Int.*, **99**, 595-604.
- Dziewonski, A.M., A.L. Hales, and E.R. Lapwood, 1975. Parametrically simple earth models consistent with geophysical data, *Phys. Earth Planet. Int.*, **10**, 12-48.
- Frazer, L.N., 1987. Synthetic seismograms using multifold path integrals - I. Theory, *Geophys. J. R. astr. Soc.*, **88**, 621-646.
- Gutowski, P.R., and E.R. Kanasevich, 1974. Velocity spectral evidence of upper mantle discontinuities, *Geophys. J. R. astron. Soc.*, **36**, 21-32.
- Haddon, R.A.W., E.S. Husebye, and D.W. King, 1977. Origins of precursors to *P'P'*, *Phys. Earth planet. Inter.*, **14**, 41-70.
- Husebye, E.S., and R. Madariaga, 1970. The origin of precursors to core waves, *Bull. Seismol. Soc. Am.*, **60**, 939-952.
- Husebye, E., R.A.W. Haddon, and D.W. King, 1977. Precursors to *P'P'* and upper mantle discontinuities, *J. Geophys.*, **43**, 535-543.
- Ingate, S.F., J. Ha, and K.J. Muirhead, 1986. Limitations on waveform modelling of long-period seismograms, *Geophys. J. R. astron. Soc.*, **86**, 57-61.
- Katsura, T., and E. Ito, 1989. The system  $Mg_2SiO_4$ - $Fe_2SiO_4$  at high pressures and temperatures: Precise determination of stabilities of olivine, modified spinel, and spinel, *J. Geophys. Res.*, **94**, 15,663-15,670.
- Kind, R., and L. P. Vinnik, 1988. The upper mantle discontinuities underneath the GRF array from *P*-to-*S* converted phases, *J. Geophys.*, **62**, 138-147.
- Langston, C. A., 1979. Structure under Mount Rainier, Washington, inferred from teleseismic body waves, *J. Geophys. Res.*, **84**, 4749-4762.
- Liu, M., and C.G. Chase, 1991. Boundary-layer model of mantle plumes with thermal and chemical diffusion and buoyancy, *Geophys. J. Int.*, **104**, 433-440.
- Liu, M., D. Yuen, W. Zhao, and S. Honda, 1991. Developments of diapiric structures in the upper mantle due to phase transitions, 1991. *Science*, **252**, 1836-1839.
- Machetel, P., and P. Weber, 1991. Intermittent layered convection in a model mantle with an endothermic phase change at 670 km, *Nature*, **350**, 55-57.

- Nakanishi, I., 1986. Seismic reflections from the upper mantle discontinuities beneath the Mid-Atlantic ridge observed by a seismic array in Hokkaido region, Japan, *Geophys. Res. Lett.*, **13**, 1458-1461.
- Nakanishi, I., 1988. Reflections of  $P'P'$  from upper mantle discontinuities beneath the Mid-Atlantic Ridge, *Geophys. J.*, **93**, 335-346.
- Neele, F., and R. Snieder, 1991. Are long-period coda wave caused by lateral heterogeneity?, *Geophys. J. Int.*, **107** 131-153.
- Nolet, G., and M. J. R. Wortel, 1989. Structure of the Upper Mantle, in *Encyclopedia of Geophysics*, pp. 775-788, ed. by James, D.E., Van Nostrand Reinhold Company, New York.
- Paulssen, H., 1985. Upper mantle converted waves beneath the NARS array, *Geophys. Res. Lett.*, **12**, 709-712.
- Paulssen, H., 1988. Evidence for a sharp 670-km discontinuity as inferred from  $P$  to  $S$  converted waves., *J. Geophys. Res.*, **93**, 10489-10500.
- Revenaugh, J., and T. Jordan, 1989. A study of mantle layering beneath the western Pacific, *J. Geophys. Res.*, **94**, 5787-5813.
- Richards, M.A., B.H. Hager, and N.H. Sleep, 1988. Dynamically supported geoid highs over hotspots: observation and theory, *J. Geophys. Res.*, **93**, 7690-7708.
- Richards, M.A., and C.W. Wicks, 1990.  $S$ - $P$  conversion from the transition zone beneath Tonga and the nature of the 670 km discontinuity, *Geophys. J. Int.*, **101**, 1-35.
- Richards, P.G., 1972. Seismic waves reflected from velocity gradient anomalies within the Earth's upper mantle, *Z. Geophys.*, **38**, 517-527.
- Sen, M.K., and L.N. Frazer, 1987. Synthetic seismograms using multifold path integrals - II. Computations, *Geophys. J. R. astr. Soc.*, **88**, 647-671.
- Shearer, P., 1990. Seismic imaging of upper mantle structure - new evidence for a 520 km discontinuity, *Nature*, **344**, 121-126.
- Shor, G.G., and D.D. Pollard, 1964. Mohole site selection studies north of Maui, *J. Geophys. Res.*, **69**, 1627-1637.
- Stammler, K., R. Kind, G.L. Kosarev, A. Plesinger, J. Horálek, L. Qiyuan and L.P. Vinnik, 1991. Broadband observations of  $PS$  conversions from the upper mantle in Eurasia, *Geophys. J. Int.*, **105**, 801-804.
- Wajeman, N., 1988. Detection of underside  $P$  reflections at mantle discontinuities by stacking broadband data, *Geophys. Res. Lett.*, **15**, 669-672.
- Walck, M. C., 1984. The  $P$ -wave upper mantle structure beneath an active spreading center: the Gulf of California, *Geophys. J. R. astron. Soc.*, **76**, 697-723.
- Whitcomb, J.H., 1973. Asymmetric  $P'P'$ : an alternative to  $P'dP'$  reflections in the uppermost mantle (0 to 110 km), *Bull. Seismol. Soc. Am.*, **63**, 133-143.
- Whitcomb, J.H., and D.L. Anderson, 1970. Reflection of  $P'P'$  seismic waves from discontinuities in the mantle, *J. Geophys. Res.*, **75**, 5713-5728.
- Yuen, D., and W.R. Peltier, 1980. Mantle plumes and thermal stability of the  $D''$  layer, *Geophys. Res. Lett.*, **7**, 625-628.

## Chapter 5

### A formalism for including amplitude data in tomographic inversions

**Summary.** Using first order ray perturbation theory and the paraxial ray approximation, a linear relation is derived that describes the first-order perturbation in amplitude due to slowness perturbations along a ray. The method allows for an efficient calculation of Frechet derivatives for body-wave amplitudes. This makes it possible to add body-wave amplitude data to large-scale tomographic inversions for slowness distributions in the Earth. Depending directly on the derivatives of the slowness field, body wave amplitudes provide a constraint on the velocity distribution independent of travel time data. For a realistic slowness distribution in the upper mantle, examples show that the region of applicability of the linearized perturbation theory to the calculation of amplitude perturbations may become very small due to the non-linear dependence of amplitudes on slowness perturbations. Travel times show a more linear dependence on slowness perturbations for perturbation of the slowness field. The examples suggest that when amplitudes and travel times are combined in an inversion for the velocity structure of the upper mantle, a non-linear travel-time inversion must be performed before amplitude data are included. This ensures that the non-linear behaviour of amplitudes due to ray shift induced by slowness perturbations are minimized.

---

This chapter has been submitted for publication as

Neele, F., J.C. Vandecar and R. Snieder, A formalism for including amplitude data in tomographic inversions. *submitted to Geophys. J. Int.*, 1992.

## 1. Introduction

Present-day tomographic studies of the Earth provide increasingly detailed images of the Earth's interior. Large data sets of teleseismic  $P$ - or  $S$ -wave travel time data reveal slowness variations on a scale of about 100 km in the upper mantle (e.g. Spakman 1991; Zielhuis 1992). However, it may be difficult to extract geodynamic parameters from these images, since important characteristics such as the thickness of a subducting slab may be difficult to resolve. Slab thickness is not resolved well by travel times alone, because varying slab thickness produces only a small travel time anomaly. This is due to the fact that for a thin slab (with accordingly sharper gradient) a ray traveling up the slab is pulled into the slab more, traversing a larger distance yet through higher velocity material than for a thick slab, producing a strong trade-off (e.g., Fischer et al. 1988). Including regional phases (Spakman 1991) or  $PP$ -wave data (van der Hilst & Engdahl 1991) can improve the resolution. A different approach is the use of additional data sets that carry independent information on the slowness distribution, such as surface wave data (Zielhuis 1992), or gravity data (Lees & Vandecar 1990). Body wave amplitudes provide another independent data set, being sensitive to the curvature of the slowness field along the ray path (Moore 1980, 1991). Small-scale numerical experiments performed by Nowack & Lutter (1988) show that a joint inversion of amplitudes and travel times can resolve sharp gradients, significantly improving on results from inverting travel times alone.

A disadvantage of including amplitudes in a tomographic inversion is the time-consuming computation of Frechet derivatives. This is much more expensive for amplitudes than for travel times (Thomson 1983), since the dependence of amplitudes on slowness perturbations is not simple. This becomes a major problem for the large data sets that are necessary to reduce the scatter in individual amplitude measurements, which are found to be considerable. Ringdal (1977) found that the variation of  $P$ -wave amplitudes among stations of a subarray of the NORSAR network (3-10 km spacing) is as large as among stations across the entire array (20-100 km spacing).

The inversion of body-wave amplitudes has been studied by Thomson (1983) and Nowack & Lutter (1988). Thomson (1983) used an extension of the ray bending equations of Julian & Gubbins (1977) to find the amplitude perturbation due to a perturbation of the slowness field. Nowack & Lutter (1988) used a Hamiltonian formulation of the dynamic ray equations of Farra & Madariaga (1987) and derived the amplitude perturbation from a perturbation of the Hamiltonian. In this paper ray perturbation theory applied to the kinematic ray tracing equations of Snieder & Sambridge (1992) is used. The paraxial ray approximation (e.g., Cerveny & Psencik 1983) is used to describe the behaviour of rays close to the reference ray. In this way the behaviour of a narrow beam of rays can be described. The formalism clearly shows how the amplitude perturbation due to a perturbation of the slowness field is obtained by calculating the behaviour of the perturbed ray beam, analogous to the methods described by Thomson (1983) and Nowack & Lutter (1988). It is shown that all computations can be performed along the reference ray, using a single differential operator with different forcing terms and boundary conditions. This allows the use of efficient matrix manipulation algorithms. The formalism can be used to efficiently calculate Frechet derivatives of amplitudes for slowness perturbations in an arbitrary reference medium. With this method it is possible to handle large amplitude data sets and thus increase the resolving power of tomographic images of the Earth. Some analytical

examples are shown for the case of a homogeneous reference model, clarifying the dependence of the amplitude on the curvature of the slowness field.

The theory is then applied to a realistic model of upper mantle slowness distribution. For several ray geometries it is shown that the domain of applicability of linearized perturbation theory to the calculation of amplitude perturbations is smaller than for travel times. For certain ray geometries this domain may become very small. This is an illustration of the absence of a stationarity principle for amplitudes. Since the relation between travel time and slowness is more linear, the examples suggest that amplitude data be included in an inversion only after a number of (non-) linear iterations with travel time data alone have converged to a model in which the ray positions are stable. Amplitude data then may be used to refine the model.

## 2. Perturbation theory for amplitudes

The aim of the theory presented here is to estimate how the amplitude along a ray is perturbed when the three-dimensionally varying slowness distribution is slightly changed. Following Farra & Madariaga (1987), the paraxial ray approximation is used here to estimate the amplitude of seismic phases. The amplitude perturbation is found by considering the behaviour of a ray beam (a reference ray and two paraxials) in the reference and perturbed medium. For the examples shown here, the ray perturbation theory of Snieder & Sambridge (1992) is used to find the perturbation of the ray position due to small variations in the slowness model. The theory is valid for a smooth medium without interfaces. A ray-centered coordinate system is used.

### Paraxial rays in reference and perturbed media

Consider a ray traced through some arbitrary slowness model  $u_0(\mathbf{r})$ . Denote this ray by  $\mathbf{r}_0(s)$ , where  $s$  is the coordinate (arc length) along the ray. If the model is perturbed slightly,  $u(\mathbf{r}) = u_0(\mathbf{r}) + \varepsilon u_1(\mathbf{r})$ , the position of the ray is also perturbed,  $\mathbf{r}(s) = \mathbf{r}_0(s) + \varepsilon \mathbf{r}_1(s)$ , to first order in  $\varepsilon$ . In the following, a ray-centered coordinate system with vector basis  $(\hat{s}, \hat{q}_1, \hat{q}_2)$  is used, where  $\hat{s}$  is directed along the ray and  $\hat{q}_1$  and  $\hat{q}_2$  are two orthogonal vectors in the plane perpendicular to the ray. On the ray  $q_1 = 0$  and  $q_2 = 0$ . The ray offset  $\mathbf{r}_1(s)$  is measured in a plane perpendicular to the ray, i.e. in the plane defined by the vectors  $\hat{q}_1$  and  $\hat{q}_2$ . Snieder & Sambridge (1992) show that to first order in  $\varepsilon$  the perturbed ray position satisfies (their equation (50))

$$\begin{aligned} \frac{d}{ds} (u_0 \dot{q}_i) - u_0 \Omega^2 q_i + u_0^2 (\hat{q}_i \hat{q}_j : \nabla \nabla (\frac{1}{u_0})) q_j - 2u_0 \Omega e_{ij} \dot{q}_j - \frac{d}{ds} (u_0 \Omega) e_{ij} q_j \\ = u_0 \hat{q}_i \cdot \nabla (\frac{u_1}{u_0}) \end{aligned} \quad (1)$$

where  $(\dot{\quad})$  indicates differentiation with respect to  $s$ ;  $\Omega$  is the rotation of the system  $(\hat{q}_1, \hat{q}_2)$  around the ray, induced by the slowness gradients along the ray;  $e_{ij}$  is the Levi-Civita tensor. See Snieder & Sambridge (1992) for a discussion of the individual terms in (1). The perturbation in the slowness field appears only in the right-hand side of the differential equation. For the purposes of this paper, the important characteristic of equation (1) is the linearity of the differential equation. The exact shape or the individual terms are not essential to the following formalism. The coupled differential equation (1) can be written in compact operator notation as

Table 1. Relative size of different quantities

Quantity	relative size	
$\mathbf{L}, \nabla \mathbf{L}$	1	depends on $u_0$
$\mathbf{F}, \nabla \mathbf{F}$	$\varepsilon$	depends on $u_1$
$\mathbf{q}$	$\varepsilon$	perturbed reference ray
$\mathbf{w}^{(i)}$	$\gamma$	paraxial ray in unperturbed medium
$\tilde{\mathbf{w}}^{(i)}$	$\varepsilon\gamma$	perturbed paraxial ray
$\Delta \mathbf{w}^{(i)}$	$\varepsilon\gamma$	perturbation in paraxial ray position

$$\mathbf{L} \mathbf{q} = \mathbf{F} \quad (2)$$

The operator  $\mathbf{F}$  and the quantity  $\mathbf{q}$  are of relative order  $\varepsilon$  (see table 1), since these depend on the perturbed slowness. Once boundary conditions for  $\mathbf{q}$  are specified, (2) can be solved numerically for any perturbation  $u_1$ . Boundary conditions are discussed below.

The amplitude in any medium follows from observing the behaviour of nearby rays. The geometrical spreading can be calculated from a surface element enclosed by a beam of rays. As noted by Farra & Madariaga (1987) and Snieder & Sambridge (1992), ray-perturbation theory can be used to estimate the position of rays close to the reference ray (see figure 1). If  $\varepsilon$  is set to zero in (2), and appropriate boundary conditions are supplied, then the solutions of

$$\mathbf{L} \mathbf{w} = 0 \quad (3)$$

are paraxial rays ( $\mathbf{w}$ ) in the unperturbed medium  $u_0(\mathbf{r})$ . A parameter  $\gamma$  is used to measure the initial (small) difference in take-off angle (ray parameter) or position of the paraxial ray, with respect to the reference ray. The quantity  $\mathbf{w}$  is therefore of relative order  $\gamma$ , as listed in table 1. Figure 1 shows a reference ray with two paraxial rays, that have a small difference in take-off angle from the reference ray. To compute the geometrical spreading in a 3-D varying medium at least two paraxial rays need to be calculated, e.g., one in each  $\hat{\mathbf{q}}_i$  direction. Denote by  $\mathbf{w}^{(1)}$  the paraxial ray that is obtained by using zero initial conditions for the  $\hat{\mathbf{q}}_2$  component and plane wave or point source initial conditions for the  $\hat{\mathbf{q}}_1$  component. Similarly, let  $\mathbf{w}^{(2)}$  be the paraxial ray obtained by using non-zero initial conditions for the  $\hat{\mathbf{q}}_2$  component only. At any point  $s$  along the reference ray, the surface element is (see figure 1)

$$dS(s) = \|\mathbf{w}^{(1)}(s) \times \mathbf{w}^{(2)}(s)\| \quad (4)$$

The ratio of surface elements at  $s = S_0$  and  $s$  is the geometrical spreading

$$J(s, S_0) = \frac{dS(S_0)}{dS(s)} \quad (5)$$

where  $S_0$  is the end point of the ray. The amplitude is then proportional to  $J^{-1/2}$ ; with  $J$  in (5) independent of  $\gamma$ , since linearity is assumed. For a point source at  $s = 0$ , the surface element  $dS(0)$  is zero and (5) cannot be used. This difficulty can be resolved by evaluating the surface element at a small distance from the source.

For a small perturbation of the medium ( $u(\mathbf{r}) = u_0(\mathbf{r}) + \varepsilon u_1(\mathbf{r})$ ) the perturbed ray position  $\mathbf{q}(s)$  is to first order in  $\varepsilon$  given by the solution of (2). To find the amplitude perturbation the

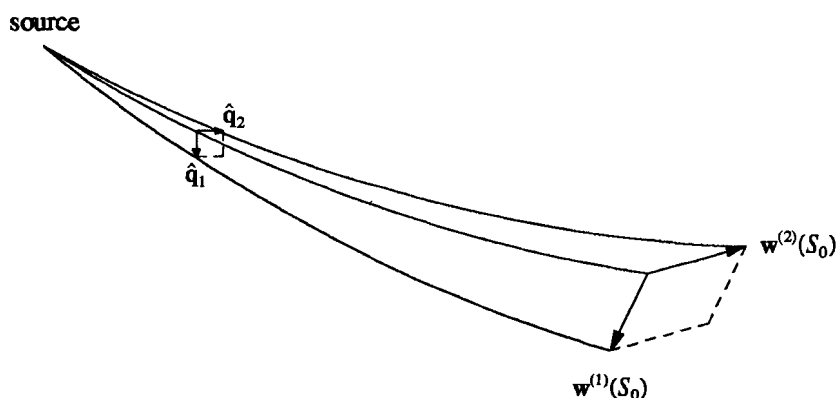


Figure 1. A ray beam consisting of a reference ray (the central ray) and two paraxial rays  $w^{(1)}$  and  $w^{(2)}$ , that have a small difference in take-off angle with the reference ray at the source in the  $\hat{q}_1$  and  $\hat{q}_2$  direction, respectively. The geometrical spreading along a ray at any point follows from the cross product  $w^{(1)} \times w^{(2)}$ .

perturbed paraxial rays must be described, to calculate the perturbed surface elements.

In order to obtain a method in which a single operator (with different forcing terms) can be used to solve all differential equations involved, equation (2) is solved along the paraxial rays  $w^{(i)}$  in the reference medium. If the appropriate boundary conditions are used (see below), this yields the paraxial rays in the perturbed medium, from which the perturbed surface elements can be calculated. Denote by  $\tilde{w}^{(i)}$  the perturbed paraxial ray, which is found by solving (2), where  $L$  and  $F$  are evaluated on  $w^{(i)}$ , the paraxial ray in the reference medium. This quantity is of relative order  $\gamma\epsilon$ , since it describes the perturbation of the paraxial ray due to a slowness perturbation (see table 1). See figure 2 for a sketch of the geometry. A first order expansion of  $L$  and  $F$  gives the operators  $L_{w^{(i)}}$  and  $F_{w^{(i)}}$  on the paraxial ray

$$L_{w^{(i)}} = L + w^{(i)} \cdot \nabla L \quad (6)$$

$$F_{w^{(i)}} = F + w^{(i)} \cdot \nabla F \quad (7)$$

In (6) and (7),  $w^{(i)} \cdot \nabla$  is proportional to the transverse gradient, as  $w^{(i)}$  lies in the  $(\hat{q}_1, \hat{q}_2)$  plane. Operators  $L$  and  $F$  without a subscript refer to the operators on the reference ray.

The perturbed paraxial  $\tilde{w}^{(i)}(s)$  is then found by solving

$$(L + w^{(i)} \cdot \nabla L)\tilde{w}^{(i)} = F + w^{(i)} \cdot \nabla F \quad (8)$$

where  $\tilde{w}^{(i)}$  is measured from  $w^{(i)}$ . The total distance of the perturbed paraxial ray  $\tilde{w}^{(i)}$  from the reference ray is then  $w^{(i)} + \tilde{w}^{(i)}$  (see figure 2). A direct measure of the difference in geometrical spreading in the perturbed and unperturbed medium, and thus of the difference in amplitude, is obtained if  $(L + w^{(i)} \cdot \nabla L)q$  is subtracted from both sides of (8). Equation (8) then becomes a differential equation for  $\tilde{w}^{(i)} - q$ . In figure 2 it is shown that this distance is equal to the difference between the distances of the paraxial ray to the reference ray and the perturbed paraxial to the perturbed ray. It is therefore a direct measure of the difference in

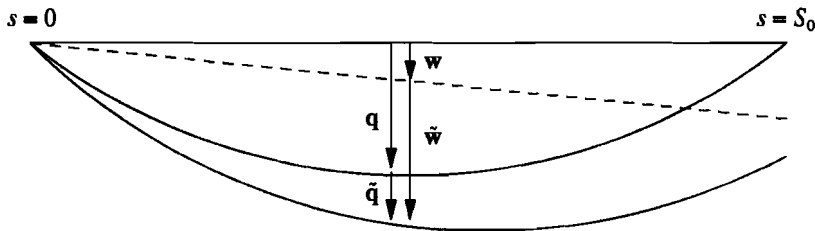


Figure 2. A schematic 2-D representation of a reference ray and paraxial ray (straight solid and dashed lines), and the perturbed ray and perturbed paraxial ray (curved lines). The source is at  $s = 0$ , the station at  $s = S_0$ . The distance of the paraxial ray in the reference medium is  $w$ ;  $\tilde{q}$  is the corresponding distance in the perturbed medium. The quantity  $\tilde{w} - q$  is equal to the difference  $\tilde{q} - w$ , and is thus a direct measure of the difference in geometrical spreading in the reference and perturbed media.

geometrical spreading between the reference and perturbed medium. Equation (8) then becomes

$$(\mathbf{L} + \mathbf{w}^{(i)} \cdot \nabla \mathbf{L})(\tilde{w}^{(i)} - q) = \mathbf{F} + \mathbf{w}^{(i)} \cdot \nabla \mathbf{F} - (\mathbf{L} + \mathbf{w}^{(i)} \cdot \nabla \mathbf{L})q \quad (9)$$

Writing  $\tilde{w}^{(i)} - q = \Delta w^{(i)}$ , and using  $\mathbf{L}q = \mathbf{F}$ , this becomes

$$\mathbf{L}\Delta w^{(i)} = \mathbf{w}^{(i)} \cdot \nabla \mathbf{F} - \mathbf{w}^{(i)} \cdot \nabla \mathbf{L}q \quad (10)$$

The quantity  $\Delta w^{(i)}$  is of order  $\gamma\epsilon$ , since it describes the perturbation in paraxial distance (relative order  $\gamma$ ) due to a slowness perturbation of order  $\epsilon$ . The term  $\mathbf{w}^{(i)} \cdot \nabla \mathbf{L}\Delta w^{(i)}$  is of an order  $\gamma$  higher than the other terms in (9); it can therefore be ignored. Note that equation (10) is linear in  $\gamma$ , as opposed to (8), which is non-linear in  $\gamma$ . Thus, not only an equation for the difference between the perturbed 2-point ray and the paraxial ray is obtained, but also the non-linearity is removed. It is shown below how the first order amplitude perturbation follows from  $\Delta w^{(i)}$ . The same difference in geometrical spreading  $\Delta w^{(i)}$  could also be obtained by ray tracing three slightly different rays through the unperturbed and perturbed media. This involves the subtraction of two large numbers, with a small difference, which is numerically unstable. Equation (10) does not have this disadvantage.

Note that the integration of (10) is performed along the reference ray. Appendix A gives the explicit form of (10). Note also, that in equations (6)-(10) the terms  $\mathbf{w}^{(i)} \cdot \nabla \mathbf{L}$  and  $\mathbf{w}^{(i)} \cdot \nabla \mathbf{F}$  do not arise from a perturbation of the slowness distribution, but from a perturbation of the ray along which the differential equation (2) is integrated. The term  $\mathbf{w}^{(i)} \cdot \nabla \mathbf{L}$  contains the variation of the reference slowness distribution  $u_0(\mathbf{r})$  in the directions  $\hat{\mathbf{q}}_1$  and  $\hat{\mathbf{q}}_2$ , whereas  $\mathbf{w}^{(i)} \cdot \nabla \mathbf{F}$  contains the variation in the slowness perturbation  $u_1(\mathbf{r})$  in these directions. Therefore, the perturbed paraxials  $\tilde{w}^{(1)}$  and  $\tilde{w}^{(2)}$  depend on the second derivative of the slowness perturbation, as  $\mathbf{w}^{(i)} \cdot \nabla \mathbf{F}$  contains terms like  $\hat{\mathbf{q}}_i \hat{\mathbf{q}}_j : \nabla \nabla \left( \frac{u_1}{u_0} \right)$ . Compare this with the dependence of the paraxials in the unperturbed medium (and through (4) and (5) also the amplitude) on the second derivative of the reference slowness, through the third term in (1). This clearly shows the dependence of the amplitude on the curvature of the slowness field (Moore, 1980, 1991).



The reason for solving (2) along the paraxial ray  $w^{(i)}$  to find the perturbed paraxial ray  $\tilde{w}^{(i)}$  is that it results in a formalism in which the amplitude perturbation can be determined using a single differential operator  $L$  with different forcing terms and boundary conditions (see below). A different approach to the calculation of the perturbed paraxials is given in Appendix B. It solves the paraxial ray equations (3) on the perturbed ray, by both expanding and perturbing the operator  $L$ . Although it appears more straightforward, the disadvantage is that for each ray and each slowness perturbation, the differential operator has to be calculated. The formalism presented here has the advantage that for each ray the operator needs to be calculated only once, and fast matrix manipulation methods can be used to calculate the amplitude perturbation due to any slowness perturbation.

### Boundary conditions

The differential equations (2), (3) and (10) need to be supplemented with boundary conditions. Two distinct cases are plane wave and point source initial conditions; mixed boundary conditions are the third possibility.

*Plane wave front.* The boundary conditions for a plane wave front at  $s = 0$  are, for paraxial rays  $w^{(1)}$  and  $w^{(2)}$  (see figure 3),

$$w^{(1)}(0) = \begin{pmatrix} \gamma \\ 0 \end{pmatrix}; \quad \dot{w}^{(1)}(0) = \mathbf{0} \quad (11a)$$

$$w^{(2)}(0) = \begin{pmatrix} 0 \\ \gamma \end{pmatrix}; \quad \dot{w}^{(2)}(0) = \mathbf{0} \quad (11b)$$

With boundary conditions (11) the starting point of the paraxial ray  $w^{(i)}$  is at a (small) distance  $\gamma$  from the reference ray in the  $\hat{q}_i$  direction, and the paraxial ray is parallel to the reference ray.

The perturbed ray  $q$  must go through the station, so the boundary conditions are

$$\dot{q}(0) = \mathbf{0}, \quad q(S_0) = \mathbf{0} \quad (12)$$

The boundary conditions for  $\Delta w^{(i)}$  follow from those for  $\tilde{w}^{(i)}$  and  $q$ , since  $\Delta w^{(i)} = \tilde{w}^{(i)} - q$ . Figure 3 illustrates the geometry for a plane wave front. At  $s = 0$ , the perturbed paraxial rays  $\tilde{w}^{(i)}$  must be at the same distance from the perturbed ray as the unperturbed paraxials from the reference ray. Only then is the quantity  $\Delta w^{(i)}(S_0)$  a direct measure of amplitude variation due to the slowness perturbation. At  $s = 0$ , the perturbed paraxial ray  $\tilde{w}^{(i)}$  is at a distance  $\gamma$  from the perturbed ray, in the direction  $\hat{q}_i$ . However,  $\tilde{w}^{(i)}$  is measured from the paraxial ray  $w^{(i)}$ , so  $\tilde{w}^{(i)}(0) = q(0)$ . With  $\Delta w^{(i)} = \tilde{w}^{(i)} - q$ , this gives  $\Delta w^{(i)}(0) = \mathbf{0}$  (see figure 3). Therefore, the boundary conditions for  $\Delta w^{(i)}$  are

$$\Delta w^{(i)}(0) = \mathbf{0}, \quad \Delta \dot{w}^{(i)} = \mathbf{0} \quad (13)$$

These conditions reflect that both unperturbed and perturbed paraxial rays start off in the same direction, and at the same distance from the reference and perturbed ray, respectively.

*Point source.* If a point source is present at  $s = 0$ , the boundary conditions are different. Paraxial rays are found by solving (2) with conditions for  $w^{(1)}$  and  $w^{(2)}$  (see figure 4 for an

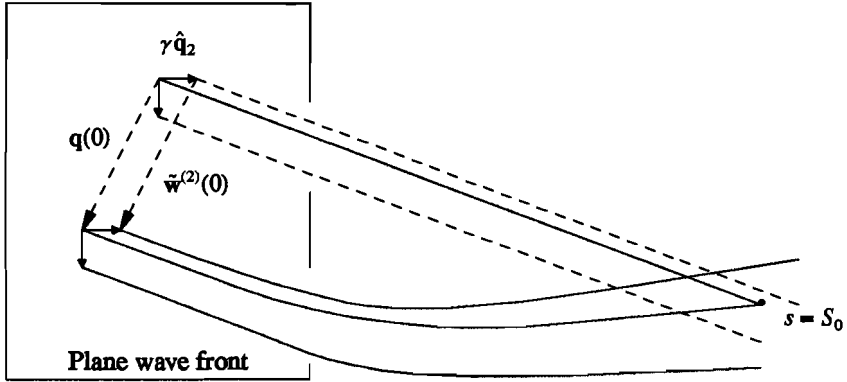


Figure 3. A schematic geometry of reference and perturbed ray beam for a plane wave front at  $s = 0$ . The station is at  $s = S_0$ . Both ray beams are perpendicular to the wave front at  $s = 0$ ; the paraxial rays in reference and perturbed media have an initial offset of  $\gamma \hat{q}_i$  from the reference ray and perturbed reference ray, respectively. Therefore, the vectors  $q(0)$  and  $\tilde{w}^{(i)}(0)$  are parallel, and the initial condition for  $\Delta w^{(i)}$ , which follows from  $q(0) - \tilde{w}^{(i)}(0)$ , is zero.

illustration of the geometry)

$$w^{(1)}(0) = \mathbf{0}, \quad \dot{w}^{(1)}(0) = \begin{pmatrix} \gamma \\ 0 \end{pmatrix} \quad (14a)$$

$$w^{(2)}(0) = \mathbf{0}, \quad \dot{w}^{(2)}(0) = \begin{pmatrix} 0 \\ \gamma \end{pmatrix} \quad (14b)$$

The perturbed ray must go through source and station, so for  $q$  the boundary conditions are the two-point ray tracing conditions

$$q(0) = \mathbf{0}, \quad q(S_0) = \mathbf{0} \quad (15)$$

As in the plane wave case, the boundary conditions for  $\Delta w^{(i)}$  can be chosen in such a way that the unperturbed and perturbed surface elements near the source are the same. The boundary for  $\Delta w^{(i)}$  are then again (13).

*Mixed boundary conditions.* In a tomographic inversion of teleseismic  $P$ -wave amplitude data for upper mantle structure, it is efficient to consider only that part of the ray path that is inside the region where the slowness is perturbed. Mixed boundary conditions have to be applied to the point of entry to the model region, to account for the geometrical spreading in the (unperturbed) Earth between source and model region. For now it is assumed that the station at  $s = S_0$ , is inside or on the boundary of the model region. The generalization to a station away from the model region is straightforward. Such a configuration can be present in a cross-borehole experiment, where both source and receiver may be outside the target zone.

The geometrical spreading in the background model between source and the point of entry to the model region can be incorporated in the boundary conditions for  $w^{(i)}$ . Denote by  $i_0$  and  $\psi$  the angle of incidence and the azimuth of the ray at  $s = 0$ , respectively. The

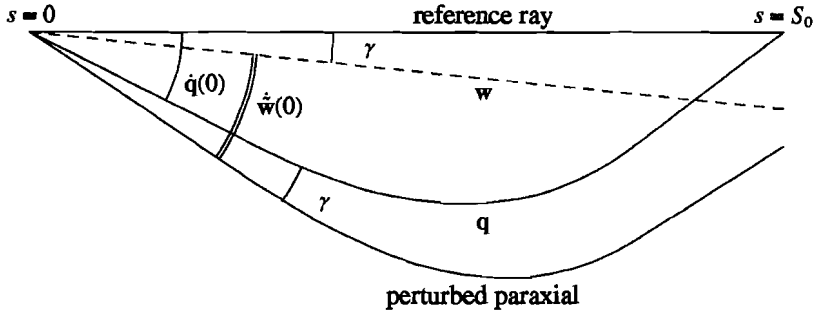


Figure 4. Schematic 2-D representation of the geometry of reference and perturbed ray beam for a point source. The paraxial ray  $w$  in the reference medium is at an angle  $\gamma$  with the reference ray; the perturbed paraxial ray is also at an angle  $\gamma$  with the perturbed ray  $q$ . As  $\Delta \dot{w}^{(i)} = \dot{q} - \dot{w}^{(i)}$ ,  $\Delta \dot{w}^{(i)}(0) = 0$ .

point  $s = 0$  is chosen at the point of entry in the region of interest. The gradients of  $i_0$  and  $\psi$  are  $\nabla i_0$  and  $\nabla \psi$ , and depend on the background model. A paraxial ray at a distance  $\gamma \hat{q}_i$  from the reference ray is then at an angle  $\gamma (\hat{q}_i \cdot \nabla i_0 + \hat{q}_i \cdot \nabla \psi)$  with the reference ray. The boundary conditions for  $w^{(i)}$  that take into account the geometrical spreading between the source and  $s = 0$  are then

$$w^{(1)}(0) = \begin{pmatrix} \gamma \\ 0 \end{pmatrix}, \quad \dot{w}^{(1)}(0) = \begin{pmatrix} \gamma \tan(\hat{q}_1 \cdot \nabla i_0 + \hat{q}_1 \cdot \nabla \psi) \\ 0 \end{pmatrix} \quad (16a)$$

$$w^{(2)}(0) = \begin{pmatrix} 0 \\ \gamma \end{pmatrix}, \quad \dot{w}^{(2)}(0) = \begin{pmatrix} 0 \\ \gamma \tan(\hat{q}_2 \cdot \nabla i_0 + \hat{q}_2 \cdot \nabla \psi) \end{pmatrix} \quad (16b)$$

Appropriate boundary conditions for  $q$  are then

$$q(S_0) = 0, \quad \dot{q}(0) = \begin{pmatrix} q_1(0) \tan(\hat{q}_1 \cdot \nabla i_0 + \hat{q}_1 \cdot \nabla \psi) \\ q_2(0) \tan(\hat{q}_2 \cdot \nabla i_0 + \hat{q}_2 \cdot \nabla \psi) \end{pmatrix} \quad (17)$$

Ignoring second order variations in  $i_0$  and  $\psi$ , the boundary conditions for  $\Delta w^{(i)}$  are again given by (13).

### Amplitudes.

With the boundary conditions as specified above, the amplitude calculations for the two cases of plane wave and point source initial conditions are similar. For the point source, (5) can not be applied directly, as  $dS(0)$  is not defined. If a small sphere around the source is considered in which the medium can be assumed homogeneous, the surface element on the sphere can be used in (5). If the source is outside the inhomogeneous region,  $dS(0)$  can be replaced by  $dS(s)$ , where  $s$  is at a some distance from the source, but outside the perturbed zone. For the unperturbed medium the geometrical spreading  $J$  is

$$J(s, S_0) = \frac{\|\mathbf{w}^{(1)}(S_0) \times \mathbf{w}^{(2)}(S_0)\|}{\|\mathbf{w}^{(1)}(s) \times \mathbf{w}^{(2)}(s)\|} \quad (18)$$

With this, amplitudes in the reference medium (i.e.  $\varepsilon = 0$ ) can be calculated. If the amplitude at  $s$  is known, then the amplitude at  $S_0$  is, using  $\mathbf{w}^{(i)}(s) = \gamma$  for the plane wave case

$$A(S_0) = A(s) J^{-\frac{1}{2}}(s, S_0) = A(s) \frac{\gamma}{\|\mathbf{w}^{(1)}(S_0) \times \mathbf{w}^{(2)}(S_0)\|^{\frac{1}{2}}} \quad (19)$$

where  $A(s)$  and  $A(S_0)$  include density and velocity information at  $s$ . The radiation pattern is assumed to be included in the amplitude at the starting point,  $A(s)$ . This result is independent of  $\gamma$ , since  $\mathbf{w}^{(i)}$  depends linearly on  $\gamma$ . This dependence is canceled in (19).

For the perturbed medium the distance between the perturbed paraxial ray and the perturbed ray is  $\Delta\mathbf{w}^{(i)} + \mathbf{w}^{(i)}$  (see figure 2). The perturbed geometrical spreading  $\tilde{J}$  then is

$$\tilde{J}(s, S_0) = \frac{\|(\Delta\mathbf{w}^{(1)}(S_0) + \mathbf{w}^{(1)}(S_0)) \times (\Delta\mathbf{w}^{(2)}(S_0) + \mathbf{w}^{(2)}(S_0))\|}{\|(\Delta\mathbf{w}^{(1)}(s) + \mathbf{w}^{(1)}(s)) \times (\Delta\mathbf{w}^{(2)}(s) + \mathbf{w}^{(2)}(s))\|} \quad (20)$$

Equations (18)-(20) are used below to derive a linear relation between model perturbations and amplitude perturbations.

*Plane wave.* With the boundary conditions (16) and (17), the expressions for the geometrical spreading are the same for plane wave and mixed boundary conditions. Equations (18) and (20) become

$$J(s, S_0) = \frac{\|\mathbf{w}^{(1)}(S_0) \times \mathbf{w}^{(2)}(S_0)\|}{\gamma^2} \quad (21a)$$

$$\tilde{J}(s, S_0) = \frac{\|(\Delta\mathbf{w}^{(1)}(S_0) + \mathbf{w}^{(1)}(S_0)) \times (\Delta\mathbf{w}^{(2)}(S_0) + \mathbf{w}^{(2)}(S_0))\|}{\gamma^2} \quad (21b)$$

as the offsets of  $\mathbf{w}^{(1)}(0)$  and  $\mathbf{w}^{(2)}(0)$  are  $\gamma\hat{\mathbf{q}}_1$  and  $\gamma\hat{\mathbf{q}}_2$ , respectively. The amplitude at  $S_0$  is  $A^{(0)}J^{-\frac{1}{2}}$ , where  $A^{(0)}$  is the amplitude on the initial wave front. The perturbed ray  $\mathbf{q}$  is at a different position on the initial wave front where the amplitude may be different from that at the position of the reference ray. This effect must be taken into account, as it is not a second order effect (Um & Dahlen, 1992). The amplitude  $A^{(0)}$  is a function of the initial point on the wave front:  $A^{(0)} = A^{(0)}(q_1(0), q_2(0))$ . The ratio  $\tilde{A}/A$  of perturbed and unperturbed amplitude then becomes

$$\begin{aligned} \frac{\tilde{A}}{A} &= \frac{A^{(0)}(q_1(0), q_2(0))}{J^{\frac{1}{2}}} \frac{J^{\frac{1}{2}}}{A^{(0)}(0, 0)} \\ &= \frac{A^{(0)}(q_1(0), q_2(0))}{A^{(0)}(0, 0)} \cdot \frac{\|\mathbf{w}^{(1)}(S_0) \times \mathbf{w}^{(2)}(S_0)\|^{\frac{1}{2}}}{\|(\Delta\mathbf{w}^{(1)}(S_0) + \mathbf{w}^{(1)}(S_0)) \times (\Delta\mathbf{w}^{(2)}(S_0) + \mathbf{w}^{(2)}(S_0))\|^{\frac{1}{2}}} \end{aligned} \quad (22)$$

To first order,  $A^{(0)}(q_1(0), q_2(0)) = A^{(0)}(0, 0) + \mathbf{q}(0) \cdot \nabla A^{(0)}$ . Furthermore, note that in (22) the term  $\Delta\mathbf{w}^{(i)}$  is of order  $\gamma\varepsilon$ , while  $\mathbf{w}^{(i)}$  is of order  $\gamma$  (see table 1). To first order in  $\varepsilon$ , (22)

becomes

$$\begin{aligned} \frac{\tilde{A}}{A} &= \left( 1 + \frac{(\mathbf{q} \cdot \nabla A^{(0)})(0,0)}{A^{(0)}(0,0)} \right) \cdot \\ &\quad \left[ 1 - \frac{\|\mathbf{w}^{(1)}(S_0) \times \Delta \mathbf{w}^{(2)}(S_0)\| + \|\Delta \mathbf{w}^{(1)}(S_0) \times \mathbf{w}^{(2)}(S_0)\|}{2\|\mathbf{w}^{(1)}(S_0) \times \mathbf{w}^{(2)}(S_0)\|} \right] \\ &= 1 + \frac{(\mathbf{q} \cdot \nabla A^{(0)})(0,0)}{A^{(0)}(0,0)} \\ &\quad - \frac{\|\mathbf{w}^{(1)}(S_0) \times \Delta \mathbf{w}^{(2)}(S_0)\| + \|\Delta \mathbf{w}^{(1)}(S_0) \times \mathbf{w}^{(2)}(S_0)\|}{2\|\mathbf{w}^{(1)}(S_0) \times \mathbf{w}^{(2)}(S_0)\|} \end{aligned} \quad (23)$$

The two last terms in the right hand side of (23) depend on  $\varepsilon$  through  $\mathbf{q}$  and  $\Delta \mathbf{w}^{(i)}$ . The latter is of order  $\gamma \varepsilon$ , and because  $\mathbf{w}^{(i)}$  is of order  $\gamma$ , equation (23) is independent of  $\gamma$  (to first order). In the limit  $\gamma \rightarrow 0$  all higher order terms in  $\gamma$  are zero, and (23) gives the amplitude perturbation to first order in  $\varepsilon$ , the slowness perturbation. The amplitude perturbation consists of two terms: one giving the change in amplitude due to the fact that the perturbed ray originates at a different position on the initial wave front than the reference ray (the term containing  $\nabla A^{(0)}$ ), and a term that is due to the change in geometrical spreading.

*Point source.* For a point source the surface elements are evaluated at a (small) distance  $\Delta s$  from the source, and the geometrical spreading in the unperturbed and perturbed media become (apart from constant factors)

$$J(\Delta s, S_0) = \frac{\|\mathbf{w}^{(1)}(S_0) \times \mathbf{w}^{(2)}(S_0)\|}{(\Delta s \gamma)^2} \quad (24a)$$

$$\tilde{J}(\Delta s, S_0) = \frac{\|(\Delta \mathbf{w}^{(1)}(S_0) + \mathbf{w}^{(1)}(S_0)) \times (\Delta \mathbf{w}^{(2)}(S_0) + \mathbf{w}^{(2)}(S_0))\|}{(\Delta s \gamma)^2} \quad (24b)$$

The ratio  $\tilde{A}/A$  of perturbed and unperturbed amplitude for a point source is then similar to (23), only in this case the change in  $A^{(0)}$  is now the change in amplitude close to the source (which contains the radiation pattern). It depends on the azimuth  $\theta$  and angle  $i$  of the ray with the vertical, so  $A^{(0)} = A^{(0)}(\theta, i)$ . To first order, the perturbations in  $\theta$  and  $i$  are proportional to  $\hat{\mathbf{q}}$ . For example, for  $\theta$  the perturbation can be written as  $\Delta \theta = (\hat{\mathbf{q}} \cdot \hat{\boldsymbol{\theta}})$ . For the point source, (23) can be written as

$$\begin{aligned} \frac{\tilde{A}}{A} &= 1 + \frac{(\hat{\mathbf{q}} \cdot \hat{\boldsymbol{\theta}}) \partial_{\theta} A^{(0)} + (\hat{\mathbf{q}} \cdot \hat{\mathbf{i}}) \partial_i A^{(0)}}{A^{(0)}(0,0)} \\ &\quad - \frac{\|\mathbf{w}^{(1)}(S_0) \times \Delta \mathbf{w}^{(2)}(S_0)\| + \|\Delta \mathbf{w}^{(1)}(S_0) \times \mathbf{w}^{(2)}(S_0)\|}{2\|\mathbf{w}^{(1)}(S_0) \times \mathbf{w}^{(2)}(S_0)\|} \end{aligned} \quad (25)$$

The fact that the terms describing the amplitude perturbation due to the change in geometrical spreading are the same in (23) and (25) (last terms in these equations) does not imply that for a given slowness perturbation the value of these terms is equal, as the paraxial rays are different for these cases. Note that  $\Delta \mathbf{w}^{(i)}$  also depends on  $\mathbf{w}^{(i)}$ , through equation (10). In the next section the case of a homogeneous reference model is treated in detail, and the

difference between (23) and (25) is shown.

Equations (23) and (25) are directly applicable for the calculation of Frechet derivatives. If the amplitude  $A_0$  for the unperturbed medium is known, then these equations give the amplitude perturbation due to the perturbation of some model element. The Frechet derivative is then simply the  $O(\varepsilon)$  terms in (23) or (25), depending on the boundary conditions of the problem.

**Example: homogeneous background model.**

As an example, the case of a homogeneous background medium is shown in detail here, for the case of a plane wave and a point source, respectively. It is assumed that the amplitude on the initial plane wave front is homogeneous; the radiation pattern of the point source is isotropic. In a homogeneous medium the reference rays are straight, and the ray-centered coordinate system is chosen not to rotate around the ray, so  $\Omega = 0$ , which in this case decouples the system (1) for  $q_1$  and  $q_2$ . The operator  $L$  is then

$$L = u_0 \frac{d^2}{ds_0^2} \quad (26)$$

*Plane wave.* In a homogeneous medium rays are straight and (for a plane wave front) parallel. The paraxial rays have the form  $\mathbf{w}^{(i)}(s) = \gamma \hat{\mathbf{q}}_i$ , which follows from (3), (26) and (11). For the homogeneous model, the perturbed ray does not have to be computed. This surprising result follows from the fact that  $\nabla L = 0$  and the term containing  $\mathbf{q}$  in (10) vanishes. Thus, to compute the amplitude, only  $\Delta \mathbf{w}^{(i)}$  needs to be found. This is most easily done using the Green's function for the operator  $L$ , which is simple for a homogeneous medium. The solution of (10) is (in operator notation) given by  $\Delta \mathbf{w}^{(i)} = G \mathbf{w}^{(i)} \cdot \nabla \mathbf{F}$ . The Green's function must satisfy the boundary conditions (13). With  $S_0$  the end point of the ray, the Green's function for a delta function slowness perturbation at  $s$  is  $G(S_0, s) = \frac{1}{u_0} (S_0 - s)$ , for  $S_0 > s$  (see figure 5 for a schematic representation of Green's function in this case). The  $\hat{\mathbf{q}}_j$  component of  $\Delta \mathbf{w}^{(i)}(S_0)$  is then given by

$$\Delta w_j^{(i)}(S_0) = G \mathbf{w}^{(i)} \cdot \nabla \mathbf{F} = \gamma \int_0^{S_0} (S_0 - s) (\hat{\mathbf{q}}_i \hat{\mathbf{q}}_j : \nabla \nabla (\frac{u_1}{u_0})) (s) ds \quad (27)$$

The unperturbed geometrical spreading (18) is  $J_0 = 1$  so that there is no amplitude decay in the unperturbed medium. The perturbed geometrical spreading (20) becomes

$$\tilde{J}(0, S_0) = \left( \frac{\Delta w_1^{(1)}(S_0)}{\gamma} + 1 \right) \left( \frac{\Delta w_2^{(2)}(S_0)}{\gamma} + 1 \right) \quad (28)$$

Note that this expression does not depend on  $\gamma$ , as  $\Delta w_i$  is proportional to  $\gamma$ . The amplitude perturbation is then from (23) and (27)

$$\frac{\tilde{A}}{A} = 1 - \frac{\varepsilon}{2} \left( \int_0^{S_0} (S_0 - s) \left[ \hat{\mathbf{q}}_1 \hat{\mathbf{q}}_1 : \nabla \nabla (\frac{u_1}{u_0}) + \hat{\mathbf{q}}_2 \hat{\mathbf{q}}_2 : \nabla \nabla (\frac{u_1}{u_0}) \right] ds \right) \quad (29)$$

to first order in  $\varepsilon$ , explicitly showing the dependence on  $\varepsilon$  of the second term in the right-

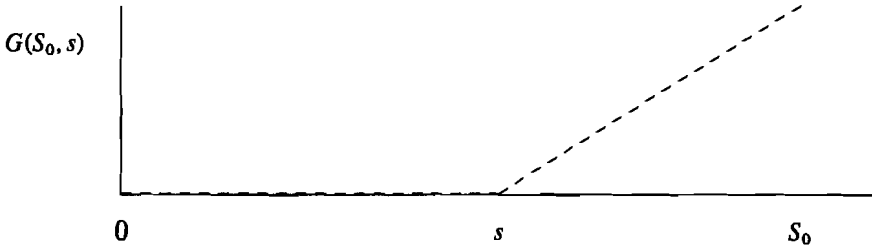


Figure 5. Illustration of Green's function  $G(S_0, s)$  for  $S_0 > s$ , for a delta function slowness perturbation at  $s$ . This  $G$  satisfies boundary conditions (13), and has a jump of  $1/u_0$  at  $s$ .

hand-side of (29). From (29) the dependence of the amplitude perturbation on the second transverse derivative of the slowness perturbation  $u_1$  is evident.

*Point source.* For a point source, the paraxials in the unperturbed medium have the form  $\mathbf{w}^{(i)}(s) = \gamma s \hat{\mathbf{q}}_i$ , which follows from (3), (26) and (14). Note that  $\gamma$  in this case is dimensionless, whereas for the plane wave it has the dimension of distance. This follows from the boundary conditions (11) and (14). As the boundary conditions for  $\Delta \mathbf{w}^{(i)}$  for a point source and plane wave are identical, the same Green's function as in the plane wave case is used to find the the solution  $\Delta \mathbf{w}^{(i)}$  for a point source

$$\Delta w_j^{(i)}(S_0) = G \mathbf{w}^{(i)} \cdot \nabla \mathbf{F} = \gamma \int_0^{S_0} (S_0 - s) s (\hat{\mathbf{q}}_i \hat{\mathbf{q}}_j : \nabla \nabla (\frac{u_1}{u_0}))(s) ds \quad (30)$$

Evaluating the surface element near the source at  $s = \Delta s$ , the unperturbed geometrical spreading  $J$  in this case is  $(S_0 / \Delta s)^2$ , so the amplitude falls off with  $1/S_0$ , as expected for a point source in a homogeneous 3-D medium. The amplitude perturbation is again given by (23) and becomes

$$\frac{\tilde{A}}{A} = 1 - \frac{\varepsilon}{2} \left( \frac{1}{S_0} \int_0^{S_0} (S_0 - s) s \left[ \hat{\mathbf{q}}_1 \hat{\mathbf{q}}_1 : \nabla \nabla (\frac{u_1}{u_0}) + \hat{\mathbf{q}}_2 \hat{\mathbf{q}}_2 : \nabla \nabla (\frac{u_1}{u_0}) \right] ds \right) \quad (31)$$

again to first order in  $\Delta \mathbf{w}^{(i)}$ . An expression of the same shape as (31) was derived by Snieder (1987) from surface wave scattering theory, describing the focusing of surface waves due to velocity perturbations.

If  $\hat{\mathbf{q}}_i \hat{\mathbf{q}}_i : \nabla \nabla (\frac{u_1}{u_0})$  is constant along the ray, the expressions take a simple form. The amplitude perturbation is then given by

$$\frac{\tilde{A}}{A} = 1 - \frac{\varepsilon}{4} S_0^2 \left( (\hat{\mathbf{q}}_1 \hat{\mathbf{q}}_1 : \nabla \nabla (\frac{u_1}{u_0})) + (\hat{\mathbf{q}}_2 \hat{\mathbf{q}}_2 : \nabla \nabla (\frac{u_1}{u_0})) \right) \quad (\text{plane wave}) \quad (32)$$

$$\frac{\tilde{A}}{A} = 1 - \frac{\varepsilon}{12} S_0^2 \left( (\hat{\mathbf{q}}_1 \hat{\mathbf{q}}_1 : \nabla \nabla (\frac{u_1}{u_0})) + (\hat{\mathbf{q}}_2 \hat{\mathbf{q}}_2 : \nabla \nabla (\frac{u_1}{u_0})) \right) \quad (\text{point source}) \quad (33)$$

### 3. Validity of the method

The theory presented here is derived using two approximations. The first is the paraxial ray approximation, associated with a scale factor  $\gamma$ , measuring the distance to the paraxial ray from the reference ray. However, the results (23) and (25) do not depend on  $\gamma$ , as this dependence is canceled to first order in the division  $J/\bar{J}$ . In the limit  $\gamma \rightarrow 0$ , the results (23) and (25) are exact as all terms of order  $\gamma^2$  that have been ignored in these equations go to zero. Therefore, the paraxial approximation does not restrict the applicability of the method presented.

The second approximation is associated with first order (linearized) ray perturbation theory, which is valid for small values of  $\varepsilon$ , where  $\varepsilon$  is a measure of the slowness perturbation. In the limit  $\varepsilon \rightarrow 0$ , the Frechet derivatives of amplitudes for slowness perturbation give the correct initial perturbation of amplitude due to a perturbation of the slowness field, as is shown in the next section. However, linearized ray perturbation theory is valid only for small values of  $\varepsilon$ . Snieder & Sambridge (1992) have derived rules for the applicability of their theory. These rules state that linearized ray perturbation theory may be applied if the number  $\varepsilon \sim u_1/u_0$  outweighs the factor  $S_0/L$ , where  $S_0$  is the ray length and  $L$  is the scale length of slowness variations. This arises from the restriction that distance between the perturbed and the reference ray may not be larger than the scale length over which the slowness perturbation varies. Although this condition must also be met here, an additional condition must be satisfied.

Equations (32) and (33) show that the amplitude depends on the second transverse derivative of the slowness perturbation, evaluated along the reference ray. The estimate of the perturbed amplitude can only be expected to be accurate if the slowness along the perturbed ray can be accurately estimated using the slowness and its derivatives along the reference ray. Generally, the second derivative of the slowness varies on the shortest length scale. Therefore, the perturbed ray may not be further from the reference ray than the scale lengths  $L_0$  and  $L_1$  over which the second derivative of either reference slowness  $u_0$  or slowness perturbation  $u_1$  changes. Thus, the additional condition is

$$|q| \ll L_0, L_1 \quad (34)$$

This can be used as an *a posteriori* check of validity of perturbation theory for the calculation of amplitudes.

For the simple case of a 2-D homogeneous background model the restriction (34) can be used to obtain a rule that can be used to verify *a priori* if perturbation theory can be applied. Consider a slowness perturbation model given by  $u_1/u_0 = \varepsilon \exp(-x^2/2L^2)$ , which describes a vertical slab-like feature. The reference ray is vertically incident, thus at constant  $x$ . Solving (1) for the perturbed ray with two-point ray tracing boundary conditions (15) yields

$$q(s) = -\frac{1}{2} \frac{x}{L^2} \frac{u_1}{u_0} s(S_0 - s) \quad (35)$$

where  $x$  is the distance of the reference ray from the centre of the slab. The largest ray perturbation is at  $s = S_0/2$ , so condition (34) implies



$$\frac{1}{8} \left( \frac{u_1}{u_0} \right) \left( \frac{x}{L} \right) \left( \frac{S_0}{L} \right)^2 \ll 1 \quad (36)$$

Except for the factor  $x/L$ , (36) is the same as equation (68) of Snieder & Sambridge (1992). For example, applying (36) to a slab with  $L = 50$  km,  $\epsilon = 0.04$  and a ray at  $x/L = 1$ , the ray length through the slab may only be on the order of 50-100 km to avoid strong non-linearities.

With this example it can also be shown that, although a necessary one, it is not a sufficient condition for first-order perturbation theory to be applicable that the last terms in the right hand side of equations (23) and (25) are small. Consider a ray that is located near the inflection point in the slowness perturbation. The second derivative is almost zero, and the last terms in (23) and (25) are small. However, the slowness gradient is near its maximum, causing large ray perturbations. These rays show in fact the strongest non-linear behaviour of the amplitude. This is also illustrated in the next section.

#### 4. Application to the upper mantle

The theory presented here is now applied to a realistic model of slowness distribution in the upper mantle. The amplitudes found with ray perturbation theory are compared with amplitudes obtained by ray tracing a narrow beam of rays. The ray tracing code used is that of Creager & Jordan (1984), adapted by Vandecar (1991). The model used is the result of a travel time inversion by Vandecar (1991). It is the result of an inversion of about 9000  $P$ -wave travel times recorded at the Washington Regional Seismic Network (WRSN). The reader is referred to Vandecar (1991) for details of the data set used and the tomographic inversion. The slowness perturbations are parameterized by splines under tension (Cline 1981). This allows for accurate ray tracing through the model. The applicability of linearized (first-order) ray perturbation theory for the calculation of amplitudes in this model is tested for a number of different ray geometries.

Figure 6 shows model UW91F3 of Vandecar (1991). The area covered by the model is the region of the WRSN. It contains slowness perturbations on various, mostly large, scale lengths with slowness perturbations ranging from about -2% to 2%. The main feature is the high-velocity region extending down from the surface to the base of the upper mantle. The background radial velocity model is Herrin (Herrin 1968), which contains no mantle discontinuities. The moho discontinuity in the Herrin model is adjusted to a smooth decrease in velocity from 8 km/s at 40 km depth to 6 km/s at the surface.

For a number of event-station combinations, amplitudes are calculated using both the linearized theory presented above and ray tracing, for different models of slowness variation in the upper mantle below the WRSN. These different models are obtained by adding some fraction of the slowness variations of UW91F3 to the background Herrin model. This can be written as

$$u = u_{Herrin} + f u_{UW91F3} \quad (37)$$

in which  $u_{Herrin}$  is the radially symmetric Herrin Earth model and  $u_{UW91F3}$  the 3-D varying slowness anomaly distribution in UW91F3. In the following experiments, the factor  $f$  is varied between -0.5 and 1.9. Frechet derivatives are calculated for the radially symmetric model ( $f = 0$  in (37)) and model UW91F3 ( $f = 1$  in (37)). Mixed boundary conditions (16) and (17) are applied at the point where the rays enter the perturbed region. A dot

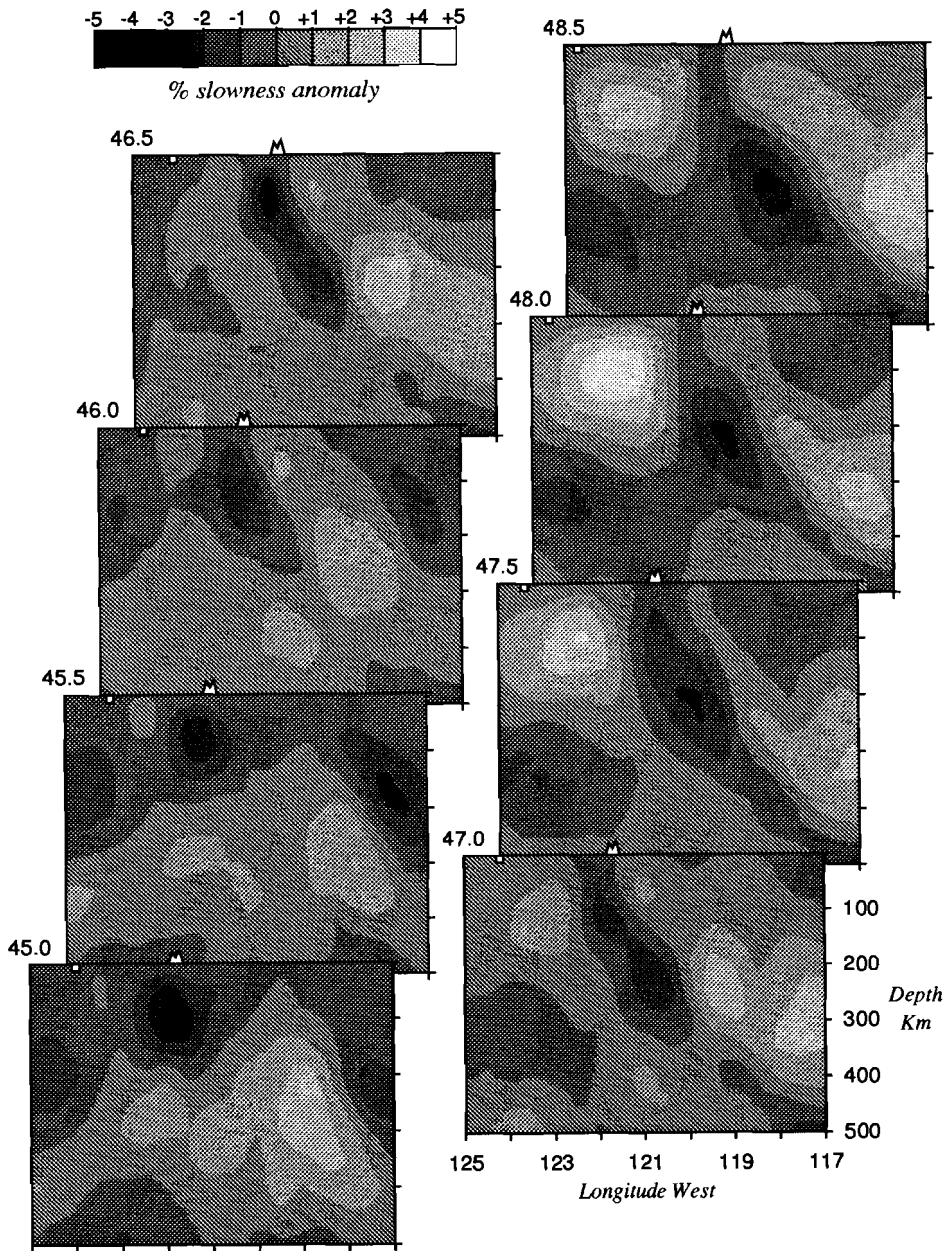
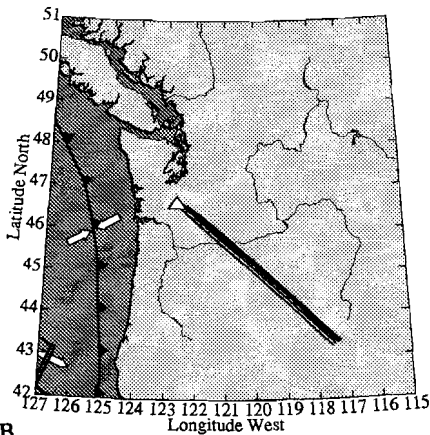


Figure 6. Model UW91F3, shown in vertical sections at constant latitude, at  $0.5^\circ$  intervals. Longitude ranges from  $117^\circ\text{W}$  to  $125^\circ\text{W}$ . Depth of the sections is 500 km. Slowness perturbations range from -4% to 4%.



A



B

Figure 7. (a), (b) Ray paths from an event at  $20^{\circ}\text{S}$ ,  $70^{\circ}\text{W}$  to WRSN station APW, shown in projections on the sagittal plane (a) and the surface (b). Heavy line is the ray path through model UW91F3, thin line represents the ray path through the radially symmetric model of Herrin [1968]. The slowness variations shown in (a) are those of model UW91F3.

product of the Frechet derivatives with the slowness perturbations (with respect to the reference model, which is either Herrin or UW91F3) along the ray then yields the linear estimate of amplitude perturbation.

The first ray path considered is shown in figure 7<sup>a</sup> and 7<sup>b</sup>, on a cross section through UW91F3; projections of the ray path on the sagittal plane and the surface are shown. The ray path is for an event at  $20^{\circ}$  south and  $70^{\circ}$  degrees west to WRSN station APW. The ray traverses a significant portion of the high velocity slab. The open line in figures 7<sup>ab</sup> is the ray path through UW91F3; the ray path through the radially symmetric Earth model (solid line) is superimposed to show the amount of ray bending induced by the model. It is clear that the ray is attracted by the high-velocity slab.

The variation of the amplitude along this ray is shown in figure 7<sup>c</sup>, as a function of the factor  $f$  in (37). Indicated by the curve connecting the squares are amplitudes obtained by tracing a beam of three slightly different rays through the model. The amplitude follows from the angle subtended by these rays at source and station. The curve connecting the

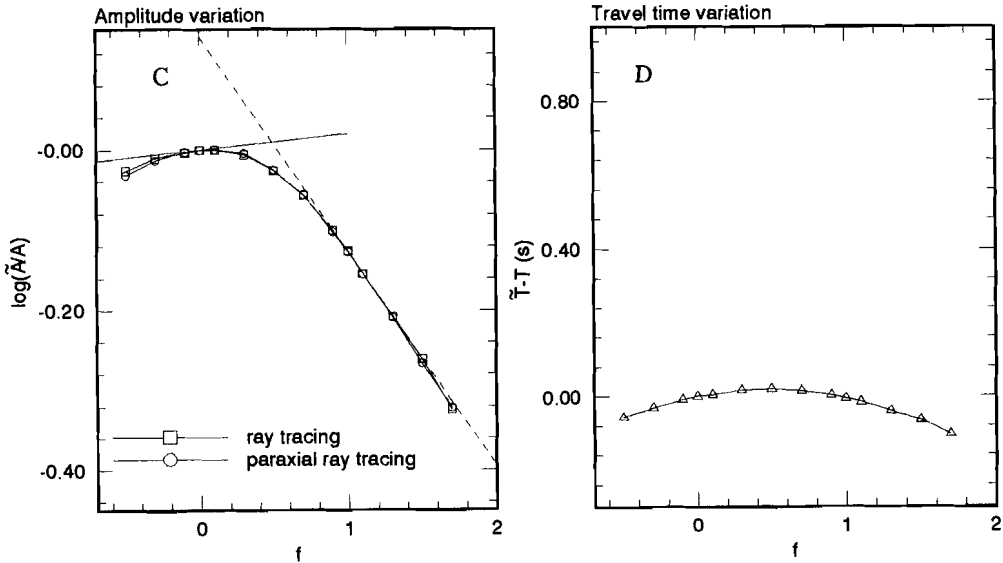
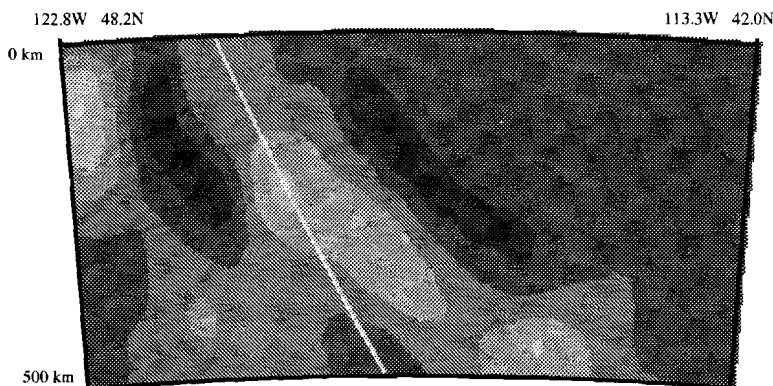


Figure 7 (continued). (c) Amplitude variations as a function of the factor  $f$  in equation (37), for the ray in (a) and (b). Squares represent amplitudes found by ray tracing, circles are amplitudes found with the paraxial approximation. The solid straight line represents the linear estimate of amplitude variations for slowness perturbations with respect to the radially symmetric model, the dashed line using UW91F3 as reference model. (d) Travel time variations for the ray in (a).

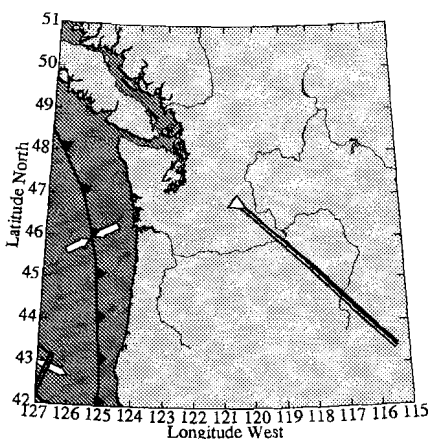
circles in figure 7<sup>c</sup> represents the amplitudes found using the ray perturbation theory presented above, using the paraxial approximation. For these two curves, the true ray paths through the different media are used. The difference between these curves is that in the latter the nearby rays are estimated using slowness information along the reference ray, implicitly assuming that the slowness derivatives are constant. The full ray tracing method involves no approximations, but has the disadvantage that it may be numerically unstable, when a tight beam of rays is traced. The solid and dashed lines, finally, represent the estimate of the amplitude perturbation using the radially symmetric model and UW91F3 as reference model, respectively.

From figure 7<sup>c</sup> two important observations can be made. First, ray perturbation theory accurately predicts the initial variation of amplitude: the straight line is tangent to the curves representing the amplitude found by actual ray tracing through the perturbed models. This shows that the Frechet derivatives correctly describe the dependence of amplitude on model parameters. Second, the amplitudes obtained by using the paraxial approximation along a true ray through the perturbed media (circles) match closely those found by full ray tracing (squares). The small differences between the curves can be removed by decreasing the angle among the three rays in the beam. The effect of this is that the volume of medium sampled by the rays becomes smaller, and thus that less averaging of slowness variations occurs. Numerical round-off errors associated with tracing a very tight bundle of rays make it impossible to reach perfect agreement between the two curves.

The non-linearity for this ray is apparent from the deviation of the curves from the linear



A



B

Figure 8. As Figure 7, now for ray paths from an event at 20°S, 70°W to WRSN station ELL.

estimate for relatively small slowness perturbation levels. The reason for this is that the ray traverses a large portion of the slab, making it sensitive to small variations in the slowness gradients of the slab, which can be varied by changing the factor  $f$ . The slowness perturbation varies strongly transversely to the ray path and relatively small changes in ray position may cause a violation of condition (38). With UW91F3 as reference model the region of linearity is much larger. The reason for this is that once the slowness perturbations are built up to the level of UW91F3, the ray is located near the centre of the slab, in a region of positive slowness curvature. This is a stable position, since increasing the slowness perturbation does not perturb the ray significantly, but merely has the effect of increasing the curvature of the slowness along the ray, and not the gradient (which causes ray shift).

For comparison, the travel time variations for this ray are shown in figure 7<sup>d</sup>, with respect to the time for the ray through the radially symmetric model. The travel times show a much more linear behaviour than the amplitudes, although there is some non-linearity.

The second example is for a ray from the same event as in figure 7 to station ELL of the

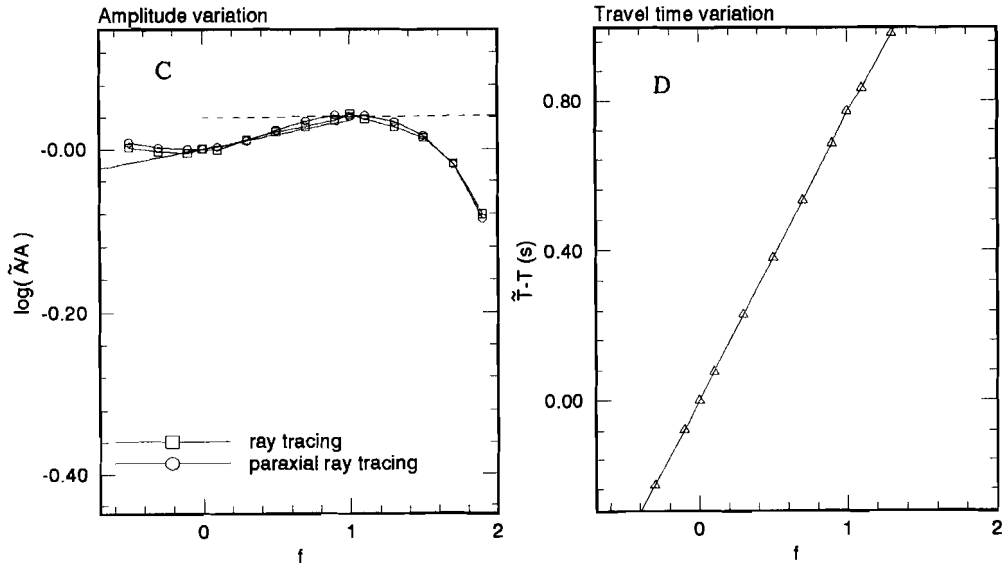
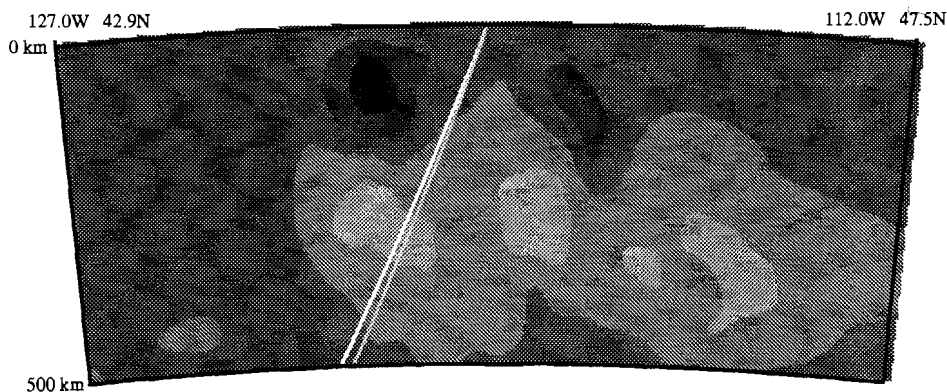


Figure 8 (continued).

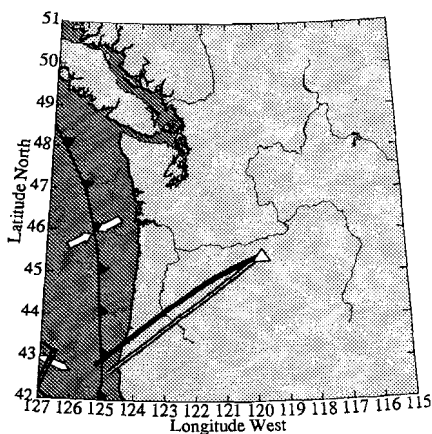
WRSN. The ray geometries are shown in figure 8<sup>ab</sup>. The amount of ray bending for this ray is small, which is reflected in the moderate amplitude variations in figure 8<sup>c</sup>. The linear estimates from ray perturbation theory again accurately predict the initial amplitude variation. However, although the amplitude variations are small, this linear estimate breaks down for slowness perturbation levels comparable to those in figure 7<sup>c</sup>. The travel time variations for the ray are almost perfectly linear (figure 8<sup>d</sup>), in contrast with the non-linear behaviour of the amplitudes.

The third example is a ray from an event located at 20°S, 179°W to WRSN station JBO (figure 9<sup>ab</sup>). Figures 9<sup>ab</sup> show that significant ray bending occurs for this ray. The amplitude variations are large for this ray (figure 9<sup>c</sup>); the change in slope of the curve in figure 9<sup>c</sup> indicates that with increasing peak slowness perturbations the ray is shifted from a focusing to a defocusing region. Ray perturbation theory accurately predicts the initial amplitude change with respect to both the radial model and UW91F3. Figure 9<sup>d</sup> shows the travel time variations, which also show non-linear behaviour for this ray.

The examples presented here illustrate the behaviour of amplitude with varying levels of slowness perturbations. An important result is that although the ray perturbation theory correctly predicts the initial amplitude change, the domain of linear behaviour of the amplitude variation may be small. This was also noted by Thomson (1983). If the perturbations are coherent along a significant part of the ray, relatively small slowness perturbations may produce strong non-linear behaviour. For example, figures 7<sup>b</sup> and 9<sup>b</sup> show that it is not warranted to use ray perturbation theory to predict the amplitude in UW91F3, using a radially symmetric Earth as reference model. For the ray geometries and the model used here, non-linearities become important when the slowness perturbations exceed about 2%. Inversion of amplitudes only, as done by, e.g., Thomson (1983), may be seriously affected by



A



B

Figure 9. As Figure 7, now for ray paths from an event at  $20^{\circ}\text{S}$ ,  $179^{\circ}\text{W}$  to WRSN station JBO.

these non-linearities. The linear domain of the travel times is much larger, which is clearest in figures 8<sup>b</sup> and 8<sup>c</sup>.

These results suggest that if travel times and amplitudes are jointly inverted for upper mantle velocity structure, amplitudes should not be included in the first iterations of a tomographic inversion. An inversion using the travel-time data, showing a more linear behaviour for small slowness perturbations, should be performed first. In this way the paths of integration for the amplitudes are closer to the true ray paths, and non-linearities are minimized. Only then is it meaningful to include amplitudes, to use those to refine the model.

## 5. Conclusions

Ray perturbation theory is applied to the calculation of amplitude perturbations due to slight perturbations in the slowness distribution along a ray. Linear relations are derived that relate slowness perturbations along a true ray path to first order perturbations in the geometrical spreading. The reference medium can be arbitrarily complex, but smooth

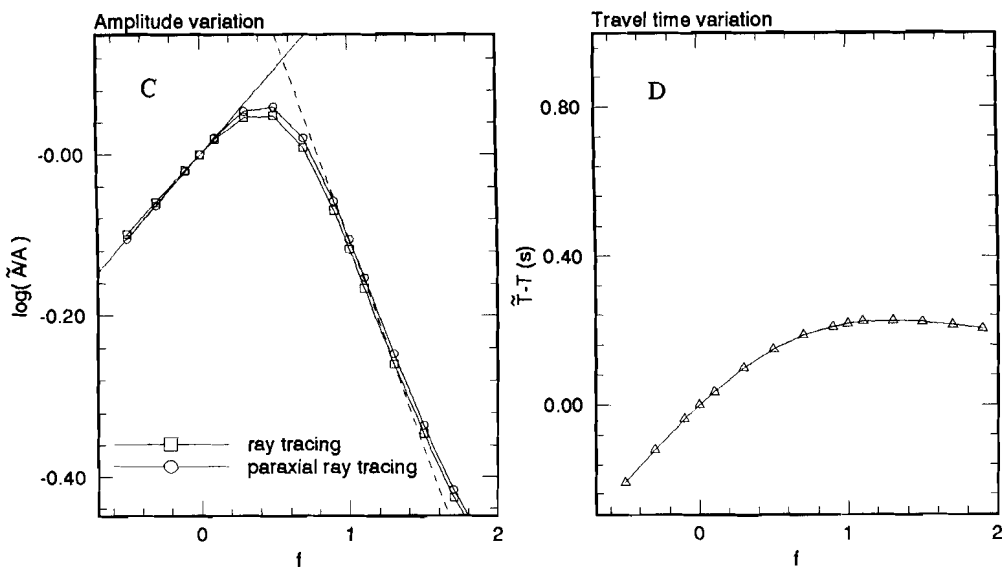


Figure 9 (continued).

without interfaces. The method presented clearly shows that the amplitude data are sensitive to the curvature of the slowness distribution and therefore add independent information to travel time data. The method allows efficient calculation of derivatives of amplitude with respect to model parameters, since all calculations can be performed along the reference ray, using the same differential operator. This makes it possible to use efficient matrix methods. The examples show that the region of linearity is smaller than the linear domain of travel times. The size of the linear region depends on the length scale and magnitude of the slowness variations in the model and the geometry of the ray path. For realistic models of slowness perturbations in the upper mantle, the examples show that it may not be warranted to use amplitudes in an inversion for slowness distribution, when the starting model is not close to the final model. The greater stability of travel times thus suggests that travel times be used in the first (non-linear) iterations and that amplitude data be included once the ray positions have become close enough to the true ray paths. In this way non-linear behaviour of the amplitude equations is minimized.

**Acknowledgement.** One of us (J.C.V.) was supported by the U.S. National Science Foundation through Grant no INT-9102113. This research was supported by the Netherlands Organization for Scientific Research (NWO) through the Pioneer project PGS 76-144,

## References

- Cerveny, V., and I. Psencik, 1983. Gaussian beams and paraxial ray approximations in three-dimensional elastic inhomogeneous media, *J. Geophys.*, 53, 1-15.  
 Cline, A.K., 1981. *FITPACK - Software package for curve and surface fitting employing splines*



- under tension*, Department of Computer Sciences, University of Texas, Austin, Texas.
- Creager, K.C. & T.H. Jordan, 1984. Slab penetration into the lower mantle, *J. Geophys. Res.*, **89**, 3031–3049.
- Farra, V., and R. Madariaga, 1987. Seismic waveform modeling in heterogeneous media by ray perturbation theory, *J. Geophys. Res.*, **92**, 2697–2712.
- Fischer, K., T.H. Jordan & K.C. Creager, 1988. Seismic constraints on the morphology of deep slabs, *J. Geophys. Res.*, **93**, 4773–4783.
- Herrin, E., 1968. Seismological tables for *P* phases, *Bull. Seismol. Soc. Am.*, **58**, 1196–1241.
- Julian, B.R., and D. Gubbins, 1977. Three-dimensional seismic ray tracing, *J. Geophys.*, **43**, 95–113.
- Lees, J.M., and J.C. Vandecar, 1990. Seismic tomography constrained by Bouguer gravity anomalies: applications in western Washington, *Pure Appl. Geophys.*, **135**, 31–52.
- Moore, B.J., 1980. Seismic ray theory for lithospheric structures with slight lateral variation, *Geophys. J. R. astron. Soc.*, **63**, 671–690.
- Moore, B.J., 1991. Seismic rays in media with slight variations in velocity, *Geophys. J. Int.*, **105**, 213–227.
- Nowack, R.L., and W.J. Lutter, 1988. Linearized rays, amplitude and inversion, *Pure Appl. Geophys.*, **128**, 401–421.
- Ringdal, F., 1977. *P*-wave amplitudes and sources of scattering in Mb-estimates, *J. Geophys.*, **43**, 611–622.
- Snieder, R., 1987. On the connection between ray theory and scattering theory for surface waves, in *Mathematical Geophysics, a survey of recent developments in seismology and geodynamics*, edited by Vlaar, N.J., G. Nolet, M.J.R. Wortel and S.A.P.L. Cloetingh, pp. 77–83, Reidel, Dordrecht, The Netherlands.
- Snieder, R., and M. Sambridge, 1992. Ray perturbation theory for travel times and ray paths in 3-D heterogeneous media, *Geophys. J. Int.*, **109**, 294–322.
- Spakman, W., 1991. Delay time tomography of the upper mantle below Europe, the Mediterranean, and Asia Minor, *Geoph. J. Int.*, **107**, 309–332.
- Thomson, C., 1983. Ray-theoretical amplitude inversion for laterally varying velocity structure below NORSAR, *Geophys. J. R. astr. Soc.*, **74**, 525–558.
- Um, J., and F.A. Dahlen, 1992. Source phase and amplitude anomalies of long-period surface waves, *Geophys. Res. Lett.*, **19**, 1575–1578.
- Vandecar, J.C., 1991. *Upper mantle structure of the Cascadia subduction zone from non-linear teleseismic travel-time inversion*, PhD thesis, University of Washington, Seattle, Washington.
- Van der Hilst, R.D., and E.R. Engdahl, 1991. On ISC PP and pP data and their use in delay-time tomography of the Caribbean region, *Geophys. J. Int.*, **106**, 169–188.
- Zielhuis, A., 1992. *S-wave velocity below Europe from delay-time and waveform inversions*, PhD thesis, University of Utrecht, Utrecht.

### Appendix A. Explicit form of equation (10).

The explicit form of the differential equations for  $\Delta \mathbf{w}^{(i)}$  are

$$\begin{aligned} \frac{d}{ds_o} (u_o \Delta w_i^{(k)}) - u_o \Omega^2 \Delta w_i^{(k)} + u_o^2 (\hat{q}_i \hat{q}_j : \nabla \nabla (\frac{1}{u_o})) \Delta w_j^{(k)} \\ - 2u_o \Omega e_{ij} \Delta w_j^{(k)} - \frac{d}{ds_o} (u_o \Omega) e_{ij} \Delta w_j^{(k)} \end{aligned} \quad (\text{A1-a})$$

$$= (\mathbf{w}^{(k)} \cdot \nabla u_o) \hat{q}_i \cdot \nabla (\frac{u_1}{u_o}) + u_o (\mathbf{w}^{(k)} \hat{q}_i : \nabla \nabla (\frac{u_1}{u_o})) \quad (\text{A1-b})$$

$$\begin{aligned} - \frac{d}{ds_o} ((\mathbf{w}^{(k)} \cdot \nabla u_o) \hat{q}_i) + (\mathbf{w}^{(k)} \cdot \nabla u_o) \Omega^2 q_i \\ - (\mathbf{w}^{(k)} \cdot \nabla u_o^2) (\hat{q}_i \hat{q}_j : \nabla \nabla (\frac{1}{u_o})) q_j \end{aligned} \quad (\text{A1-c})$$

$$\begin{aligned} - u_o^2 ( [\mathbf{w}^{(k)} \hat{q}_i \hat{q}_j : \nabla \nabla \nabla (\frac{1}{u_o}) ] ) q_j + 2(\mathbf{w}^{(k)} \cdot \nabla u_o) \Omega e_{ij} \hat{q}_j \\ + \frac{d}{ds_o} ( (\mathbf{w}^{(k)} \cdot \nabla u_o) \Omega ) e_{ij} q_j \end{aligned} \quad (\text{A1-d})$$

where (A1-a) is the operator  $L$  on the reference ray; (A1-b) is the term  $\mathbf{w}^{(k)} \cdot \nabla F$ , and (A1-c,d) represent the gradient of the operator  $L$ :  $\mathbf{w}^{(k)} \cdot \nabla L$ . All derivatives are evaluated on the reference ray. In these equations  $\Delta w_i^{(k)}$  is the  $\hat{q}_i$ -coordinate of the vector  $\Delta \mathbf{w}^{(k)}$ . The superscript  $(k)$  means that nonzero boundary conditions have been used for the  $\hat{q}_k$  component of the paraxial ray  $\mathbf{w}^{(k)}$ . This differential equation for  $\Delta \mathbf{w}^{(k)}$  can be written in matrix form; it results in a matrix that is empty, except for 9 diagonals. This matrix equation can be rewritten in the form of an equation involving a banded matrix with 6 diagonals. Efficient routines exist to solve such equations.

### Appendix B. A different approach to find the amplitude perturbation.

In the method presented in this paper the amplitude perturbation is found by computing the perturbed paraxial ray, or more precisely, finding the distance between the perturbed reference and paraxial rays. Comparing this with the distance between the reference and paraxial rays in the unperturbed medium gives the amplitude perturbation. Another possibility is to solve for the paraxial ray  $\tilde{q}$  to the perturbed reference ray  $q$  (see figure 2) in the perturbed medium, i.e., solve the differential equation

$$L_q \tilde{q}^{(i)} = 0 \quad (\text{B1})$$

where subscript  $q$  indicates that the operator is evaluated on the perturbed ray  $q$ , and superscript  $(i)$  indicates that nonzero boundary conditions are used for the  $\hat{q}_i$  component of  $\tilde{q}(0)$ . Note that (B1) is of the form (3).

As before, the operator  $L_q$  can be related to the operator on the reference ray. The operator  $L_q$  differs in two ways from the operator  $L$ . Firstly,  $L_q$  does not depend on  $u_o$ , but on  $u = u_o + \varepsilon u_1$ . Secondly,  $L_q$  is evaluated on the perturbed ray, not on the reference ray. To first order in  $\varepsilon$ , the operator  $L_q$  becomes

$$\mathbf{L}_q = \mathbf{L}(u_0) + \mathbf{L}(u_1) + \mathbf{q} \cdot \nabla \mathbf{L}(u_0) \quad (\text{B2})$$

The term  $\mathbf{q} \cdot \nabla \mathbf{L}(u_1)$  is of the order  $\varepsilon^2$ , whereas the other terms in (B2) are of order 1 ( $\mathbf{L}(u_0)$ ) and  $\varepsilon$  ( $\mathbf{L}(u_1)$  and  $\mathbf{q} \cdot \nabla \mathbf{L}(u_0)$ ), and has therefore been ignored.

With  $\mathbf{L}$  given in (1), (C1) becomes

$$\begin{aligned} \mathbf{L}_q \tilde{\mathbf{q}} = & \frac{d}{ds_0} (u_0 \dot{\tilde{q}}_i) - u_0 \Omega^2 \tilde{q}_i + u_0^2 (\hat{\mathbf{q}}_i \hat{\mathbf{q}}_j : \nabla \nabla (\frac{1}{u_0})) \tilde{q}_j - 2u_0 \Omega e_{ij} \dot{\tilde{q}}_j \\ & - \frac{d}{ds_0} (u_0 \Omega) e_{ij} \tilde{q}_j \end{aligned} \quad (\text{B3-a})$$

$$\begin{aligned} + & \frac{d}{ds_0} (u_1 \dot{\tilde{q}}_i) - u_1 \Omega^2 \tilde{q}_i + 2u_0 u_1 (\hat{\mathbf{q}}_i \hat{\mathbf{q}}_j : \nabla \nabla (\frac{1}{u_0})) \tilde{q}_j \\ & - u_0^2 (\hat{\mathbf{q}}_i \hat{\mathbf{q}}_j : \nabla \nabla (\frac{u_1}{u_0^2})) \tilde{q}_j - 2u_1 \Omega e_{ij} \dot{\tilde{q}}_j - \frac{d}{ds_0} (u_1 \Omega) e_{ij} \tilde{q}_j \end{aligned} \quad (\text{B3-b})$$

$$\begin{aligned} + & \frac{d}{ds_0} ((\mathbf{q} \cdot \nabla u_0) \dot{\tilde{q}}_i) - (\mathbf{q} \cdot \nabla u_0) \Omega^2 \tilde{q}_i \\ & + (\mathbf{q} \cdot \nabla u_0^2) (\hat{\mathbf{q}}_i \hat{\mathbf{q}}_j : \nabla \nabla (\frac{1}{u_0})) \tilde{q}_j + u_0^2 ([\mathbf{q} \hat{\mathbf{q}}_i \hat{\mathbf{q}}_j : \nabla \nabla \nabla (\frac{1}{u_0})]) \tilde{q}_j \\ & - 2(\mathbf{q} \cdot \nabla u_0) \Omega e_{ij} \dot{\tilde{q}}_j - \frac{d}{ds_0} ((\mathbf{q} \cdot \nabla u_0) \Omega) e_{ij} \tilde{q}_j \end{aligned} \quad (\text{B3-c})$$

in which (B3-a) is  $\mathbf{L}(u_0)$ , (B3-b) represents  $\mathbf{L}(u_1)$ , and (B3-c) gives  $\mathbf{q} \cdot \nabla \mathbf{L}(u_0)$ .

Using (B3), two paraxial rays  $\tilde{\mathbf{q}}^{(1)}$  and  $\tilde{\mathbf{q}}^{(2)}$  are found. If identical boundary conditions are used for  $\mathbf{w}^{(i)}$  and  $\tilde{\mathbf{q}}^{(i)}$ , the ratio of  $J$  and  $\tilde{J}$  becomes

$$\frac{\tilde{J}}{J} = \frac{\|\tilde{\mathbf{q}}^{(1)}(S_0) \times \tilde{\mathbf{q}}^{(2)}(S_0)\|}{\|\mathbf{w}^{(1)}(S_0) \times \mathbf{w}^{(2)}(S_0)\|} \quad (\text{B4})$$

Note that again this result is independent of the paraxial parameter  $\gamma$ , as both  $\tilde{\mathbf{q}}^{(i)}$  and  $\mathbf{w}^{(i)}$  are proportional to  $\gamma$ . The result (B4) depends only on  $\varepsilon$ .

The disadvantage of this method is that to compute Frechet derivatives of amplitude for slowness perturbations, (B2) has to be recomputed for every model element. Therefore, fast matrix manipulation routines cannot be used.

## Chapter 6

### The use of P-wave amplitude data in a joint inversion with travel times for upper-mantle velocity structure

**Abstract.** A joint inversion of synthetic *P*-wave travel-time and geometrical spreading amplitude data is performed to test whether the amplitude data can be used to increase the resolution of inversions using travel times alone. Geometrical spreading amplitudes depend on the curvature of the slowness field, and may thus help to resolve sharp gradients. Travel times and amplitudes of short-period *P*-waves observed at the Washington Regional Seismic Network are jointly inverted for upper mantle velocity structure below the array. The processing of the observed *P*-wave amplitudes is shown in detail, and it is shown that it is important to correct the raw data for station statics, as these dominate the observed amplitude signal. A robust averaging procedure is used to identify and remove outliers from the data set. The data comprise 8697 travel time and 4255 amplitude measurements. The travel-time data alone are used to obtain a reference model for the amplitude data. In this way the non-linear behaviour of the amplitudes is minimized. The results show that the amplitude data induce small, short-scale slowness perturbations to the starting model, producing an amplitude misfit reduction to an arbitrary degree, depending on the regularization.

Inversions with synthetic data are performed to explain this result. An idealized model of a subducting slab is used to generate synthetic data sets. The model resulting from a joint inversion produces an amplitude misfit reduction of 77%, while the travel-time misfit is unaffected. In agreement with the results for the real data sets, this is achieved by making small adjustments

---

This chapter has been accepted for publication as

Neele, F., J.C. VanDecar and R. Snieder, The use of P-wave amplitude data in a joint inversion with travel times for upper-mantle velocity structure, *J. Geophys. Res.*, *in press*, 1993.

throughout the model, which do not alter the overall slowness variations. The amplitude data do not improve the resolution of the sharp gradients present in the synthetic slab model. These results can be explained by both an extremely strong sensitivity of the geometrical spreading amplitude to slowness perturbations along the ray path, and to the distribution of the amplitude data set, which is not complete enough to induce more than incoherent changes to the starting model. The strong sensitivity is due to the combination of long teleseismic ray paths and small-scale (about 30 km) slowness variations allowed by the model parameterization. With the present amplitude data set the curvature is changed on the scale of the node spacing in the model, whereas the overall slowness variations on a scale length of several times the node spacing are unaffected. The regularization has little effect on these small-scale changes, as these changes increase the total model roughness by only 1%.

It is concluded that with the amplitude data set available, amplitude data do not improve the resolving power of travel-time inversions for upper-mantle velocity structure. This is due to the number of data required to both decrease the observed variance in the raw amplitudes to the variance in travel-time data and to obtain a good coverage of the model. The application of *P*-wave amplitude data as a validation tool is suggested for models obtained with, for example, travel-time tomography. Applications to both synthetic and observed data sets are shown. Anelastic damping is ignored in the inversions. Using linear relations between velocity, temperature and *Q*, it is shown that the effect of anelastic attenuation is on the order of 10-15% of that of geometrical spreading through the upper mantle.

## 1. Introduction

In this paper the use of short-period *P*-wave amplitudes in tomographic inversions for upper mantle velocity structure is studied. If interpreted as the result of elastic focusing and defocusing, *P*-wave amplitudes yield information on the derivatives of the velocity field (e.g., Moore, 1980), and, when combined with travel times, provide independent information on velocity variations in the Earth. In recent years, travel-time tomography has grown into a widely used tool to image the Earth's upper mantle (e.g., Zhou & Clayton, 1990; Spakman, 1991; VanDecar, 1991; van der Hilst & Engdahl, 1991). Data sets of larger size and higher quality provide increasingly detailed images of velocity variations. However, sharp velocity gradients may be difficult to resolve using travel-time data alone; such resolution is desirable when geodynamic inferences are to be made from the tomographic images. Depending directly on the derivatives of the slowness field, body wave amplitudes may help constrain sharp gradients.

*P*-wave amplitudes have received considerable attention in the past. Berteussen et al. (1975) and Flatte & Wu (1988) interpreted short-period *P*-wave amplitude variations observed across the NORSAR array in terms of scattering in a random medium. The greater part of observed fluctuations could be explained on a statistical basis. On a much larger scale, *P*-wave amplitude variations across the United States have been attributed to large-scale variations of *Q* in the upper mantle (e.g. Lay & Helmberger, 1983; Butler, 1984). These authors observed a correlation between low amplitudes and late arrivals.

Such a correlation is expected, if it is assumed that regions of low  $Q$  are caused by a generally higher temperature. However, Butler (1983) observed that early arrivals in the western United States generally have low amplitudes. This can be explained by a predominance of elastic (de)focusing over anelastic effects, regions of high velocity being associated with defocusing and hence low amplitudes.

A number of authors have attempted to use  $P$ -wave amplitudes to construct a model of velocity or  $Q$  structure. In attenuation tomography observed amplitude variations are attributed entirely to the anelasticity of the Earth (e.g., Evans & Zucca, 1988; Ho-Liu et al., 1988). Elastic (de)focusing is ignored in these studies. Haddon & Husebye (1978) inverted observed  $P$ -wave travel times at NORSAR for upper mantle velocity structure and found that the observed amplitude fluctuations compared reasonably well with the variations in geometrical spreading predicted by their model. This suggests that elastic (de)focusing in the upper mantle below the array dominates the  $P$ -wave amplitude variations. An inversion of these amplitude data was performed by Thomson (1983), using ray theory to calculate geometrical spreading in a heterogeneous medium. The models thus obtained compared only moderately well with those obtained from a  $P$ -wave travel-time inversion by Thomson & Gubbins (1982). Thomson (1983) attributed this to the amplitude data quality or validity of ray theory. Moderate amplitude data quality was caused by large scatter in the individual measurements. This was already noted by Ringdal (1977), who found that  $P$ -wave amplitude variations between stations of a subarray at NORSAR are as large as variations among stations across the whole array. This reflects the strong sensitivity of amplitudes to changes in the ray path and makes it necessary to use large data sets to reduce this scatter.

The question of applicability of linearized ray theory to the inversion of  $P$ -wave amplitudes is addressed by Nowack & Lutter (1988), who performed joint inversions of synthetic travel-time and amplitude data. For a model setup resembling a cross-borehole experiment, the joint inversions showed that amplitudes help constrain sharp gradients, which may be difficult to resolve with travel times alone. For realistic models of slowness distribution in the upper mantle and ray paths of teleseismic  $P$ -waves, Neele et al. (1992) found that the linearized theory may break down for small slowness perturbations from the reference model. Although this depends on the ray geometry and the length scales of slowness perturbations in the model, this may be important when the ray paths in the reference model are not close to the true ray paths.

In the present study the usefulness of  $P$ -wave amplitudes in large-scale tomographic inversions for upper mantle structure is studied. In section 2 the data selection and processing is described. The data used here are short-period  $P$ -wave amplitudes observed by the Washington Regional Seismic Network (WRSN). Since a careful data analysis is critical in obtaining a good data set, the procedure is described in detail. Section 3 describes the joint inversion of real travel-time and amplitude data sets, using the results of a travel-time inversion of VanDecar (1991) as a starting model. The amplitude data are interpreted as the result of geometrical spreading in the upper mantle below the array. The effect of anelastic damping is shown to be only 10-15% of that of elastic (de)focusing. The results show that the amplitude data may be fit to an arbitrary degree, if the smoothness constraints on the model are sufficiently lowered. This either points to a low variance in the data, which is unlikely, or to a strong sensitivity of geometrical spreading amplitudes to slowness perturbations along the ray path. The structures that explain the amplitude data seem to be small, short-scale slowness fluctuations, that do not affect the overall slowness patterns. To

explain the results from section 3, joint inversions with synthetic data are described in section 4. With a known slowness model, it can be tested if amplitude data increase the resolution of the inversion. The synthetic model is an idealized model of a subducting slab. In agreement with the results for the real data sets, the examples show that small changes in model parameters suffice to explain the data set, even if the data are not related to the model, i.e. if the data are derived from a model that is not similar or close to the reference model. This is a result of the strong sensitivity of geometrical spreading amplitudes on the velocity derivative distribution in the model.

The results suggest that for the data set used here amplitude data are not useful in inverse modeling for upper-mantle velocity structure, but may be used as a validation tool for models obtained with other data sets (e.g. travel times). This application of amplitude data is shown in section 5 for models from different steps in a non-linear inversion of the synthetic travel-time data. The validation method is also applied to the real amplitude data, and models obtained in a non-linear travel-time inversion described by VanDecar (1991).

## 2. Data selection and processing

The data selection and processing is presented in this section. The processing is shown in some detail here, since it is critical in obtaining a reliable amplitude data set.

*Event selection.* The data used in this study are from the Washington Regional Seismic Network Network (WRSN) in the North Western United States, shown in figure 1. This network is situated favourably for obtaining a good azimuthal coverage of events. The array has been operating digitally since 1980, with on average 120 short-period channels recording teleseismic events. The band-width of the instruments is narrow, centered on 1 second. Data have been extracted from the 10-year interval 1980-1990.

Stringent criteria for the amplitude data are used to obtain a high quality data set. Events are selected to have a high signal to noise ratio and an impulsive *P*-wave onset. Furthermore, the events are required to be grouped in tight clusters, covering as well as possible the entire azimuth range. This makes it possible to invert for average responses and locate outliers. 130 events are selected, arranged in 28 clusters, yielding 7398 good quality *P*-wave amplitudes. Epicentral distances are larger than  $30^\circ$ , so upper mantle triplications do not influence the amplitudes. The distribution of the clusters with azimuth and distance is shown in figure 2.

*Amplitude measurement.* The *P*-wave amplitude used here is the baseline to first peak amplitude. Amplitudes measured later in the signal, e.g. peak-to-peak amplitude (used by, e.g., Berteussen et al., 1975; Butler & Ruff, 1980) or the energy in a time window containing the *P*-wave (Haddon & Husebye, 1978) may be affected by shallow receiver or source structure (peg-leg multiples following directly after *P*). The baseline is estimated by averaging over a 10 second noise window prior to the *P*-wave. This does not cause problems, as the noise level is generally low. The variance in the data can be estimated by measuring the variation in amplitude observed at a given station from closely spaced events. It is estimated to be about a third order of magnitude, or  $0.3 \log(A)$ , which yields a S/N ratio of about 2-3 for the real data set.

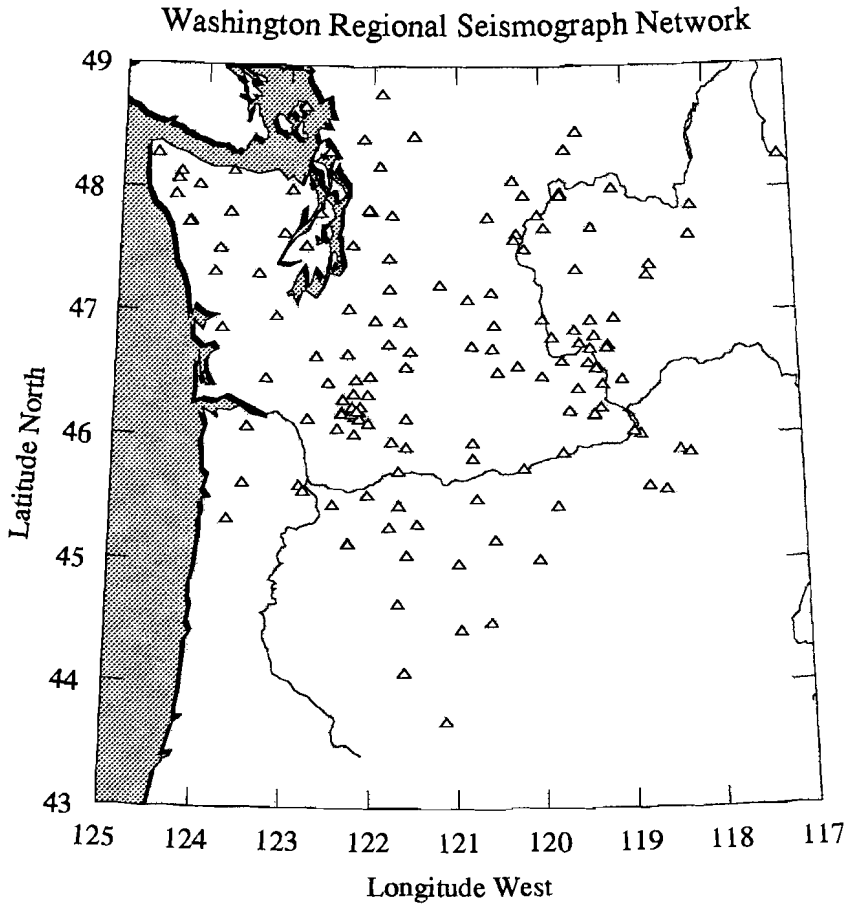


Figure 1. Configuration of the WRSN array in the north-western part of the United States.

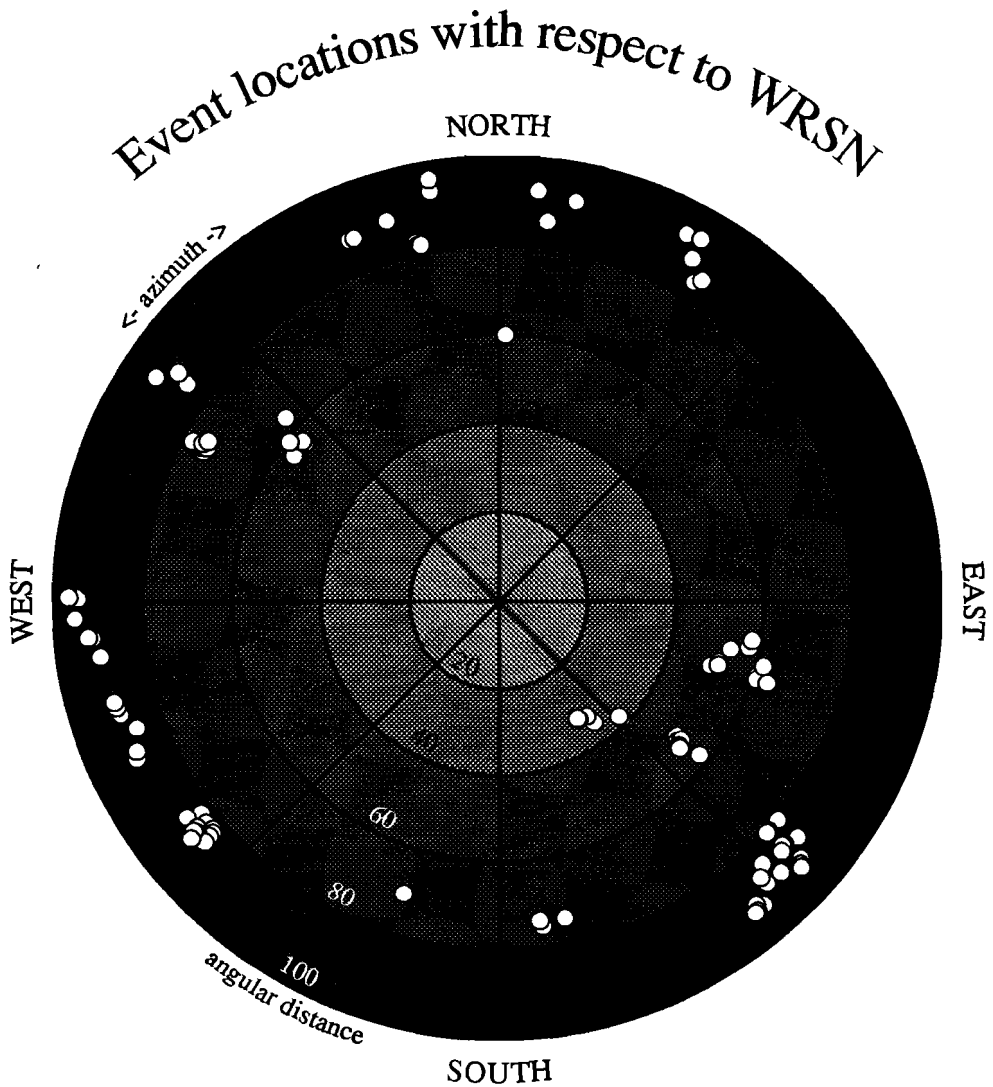
*Site effects.* Amplitudes have a large scatter, as has been noted by several authors (e.g. Ringdal, 1977). Although amplitudes tend to exhibit a near normal distribution in the log domain (Ringdal et al., 1972), outliers are often present. It is important to identify and remove these outliers.

Before outliers can be found in the data set, source statics have to be removed from the raw amplitudes. Ignoring random noise, the amplitudes can be viewed as the product of source and station statics, and the propagation effects along the path. Symbolically,

$$A_k = S_i P_k R_j \quad (1)$$

where  $A_k$  is the amplitude for ray  $k$ ,  $S_i$  is the source effect for event  $i$ ,  $R_j$  is the receiver (site) effect for station  $j$  and  $P_k$  is the effect of propagation through the region of interest, which in this case is the upper mantle below WRSN. Note, in (1),  $i$  and  $j$  determine  $k$  completely. It is assumed here that any effects from propagation near the source and through the upper mantle near the source and the lower mantle are absorbed by the source static.





*Figure 2.* Location of the events used in this study. The events are grouped in tight clusters; this makes it possible to average

The latter assumption is reasonable, since the Fresnel zone of  $P$ -waves in the lower mantle is considerable (Nolet, 1992) and a stochastic analysis of travel-time data suggests that slowness variations in the lower mantle are dominated by large-scale structure (Gudmundsson et al., 1990). The source static also contains the event magnitude and radiation pattern. For teleseismic events the array is sufficiently narrow for the radiation pattern to be considered constant across the array. All local effects are assumed to be contained in the station

static  $R$ , such as the effect of local shallow geology, instrument response and gain setting. It is further assumed that the station static does not depend on azimuth. For regional events it may be important to consider the azimuth dependence of  $R$ .

Taking the log (base 10) of equation (1) it becomes

$$\log(A_k) = \log(S_i) + \log(P_k) + \log(R_j) \quad (2)$$

If there is no a priori information about Earth structure (and thus about the path effect), the logical step to take would be to average observed amplitudes from all directions and assume that the path effects average out. Equation (2) can then be used to invert for site effect  $R$  and source static  $S$ , with  $P$  considered noise. However, in this case good estimates of the geometrical spreading through the upper mantle below the array are available, for the moment ignoring the effects of anelastic attenuation. Travel-time tomographic inversions for slowness distribution in the upper mantle below the WRSN region by VanDecar (1991) show a high-velocity slab subducting to the east (figure 3). Rays traveling up the slab are expected to be more affected than those traveling perpendicular to the slab. Therefore, it is reasonable to assume that for events from westerly azimuths amplitude variations due to 3-D velocity structure are relatively small. This a priori knowledge is used to estimate the site response by assuming that for events with back azimuths between  $225^\circ$  and  $315^\circ$  the path effects average out and that station statics dominate the amplitude signal.

Equation (2) is solved using a robust inversion technique of iterative residual down-weighting (Huber, 1981; Egbert & Booker, 1986). The technique iteratively down-weights all measurements that lie further than a specified number of standard deviations from the mean; we used 1.5 standard deviations (see VanDecar, 1991). In this way, outliers have a minimized influence on the station statics. For each station, data from all azimuths are then corrected using these site statics. Figure 4 shows the station statics obtained from the data.

*Identification of outliers, composite rays.* The events are arranged in tight clusters, which makes it possible to average the amplitudes observed at a given station over the events in a cluster, and thus to identify outliers in the data. This is done by grouping closely spaced clusters together into groups of 5 or more events. Corrected for station statics, the data can now be written as

$$\log(A_k) = \log(S_i) + \log(P_k) \quad (3)$$

Applying again the robust inversion procedure, the source terms  $S$  are estimated. This can be done most reliably if many stations recorded the same event. For each group of events, stations that recorded less than 4 events are discarded, as in that case outliers can not be identified. This reduces the data set to 5060 observations; 391 of these are outliers and are discarded. For each station, the remaining amplitudes are then averaged over the events in the single clusters. Averaged amplitudes with a standard deviation of more than 0.4 (log amplitude) are also removed. The data set is thus reduced to 1260 composite rays. Through this averaging process, the variance in the data is reduced by a factor of about 2, and the S/N ratio increased to about 5.

*Comparison of observed and predicted amplitude patterns.* The amplitude variations induced by the subducting slab are expected to be strongest for rays that travel parallel to the slab. This is the case for events with easterly and south-easterly back azimuths. For an

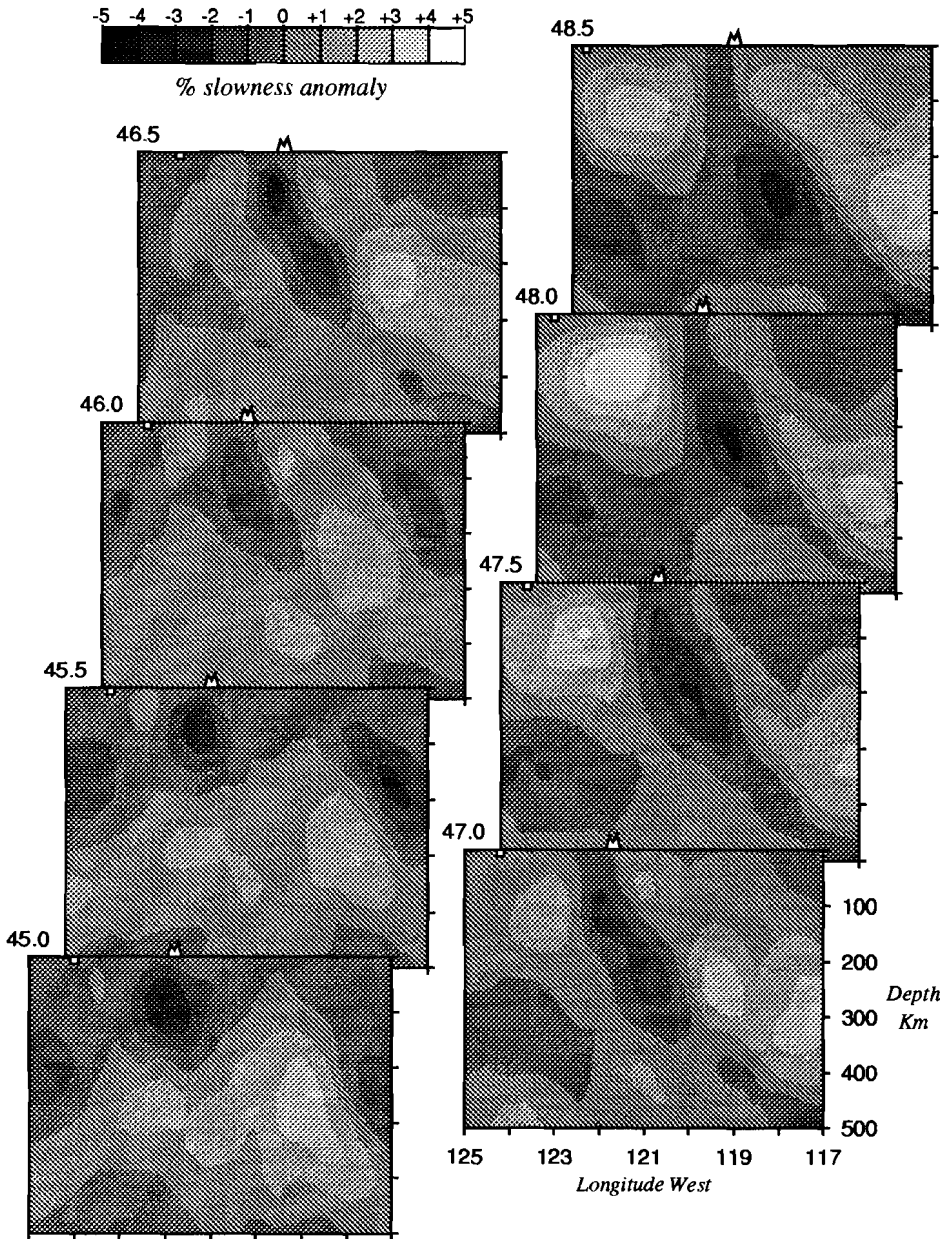


Figure 3. Slowness distributions of model UW91F3, obtained by Vandecar (1991) in a non-linear tomographic inversion of observed travel-time data. Slowness variations range from about -4% to 4%.

## Station statics from events with westerly azimuths

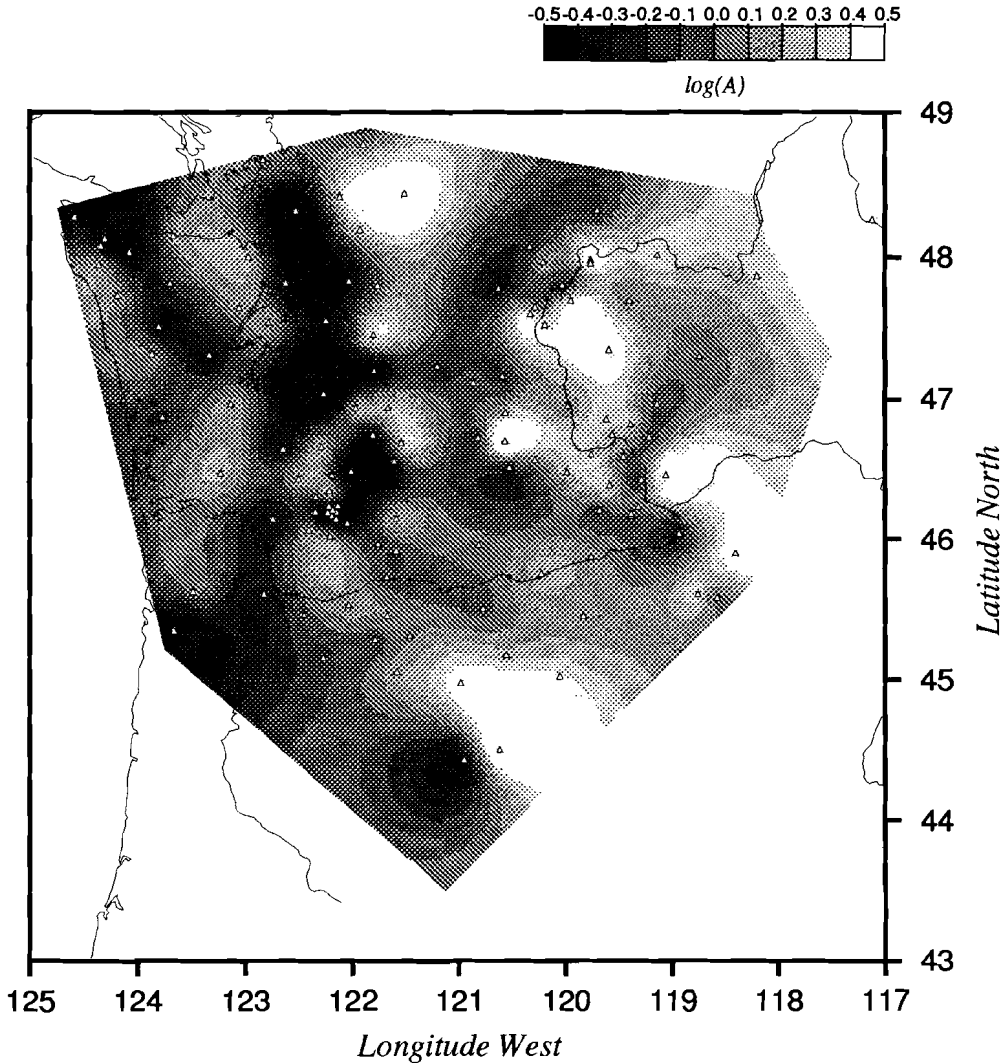
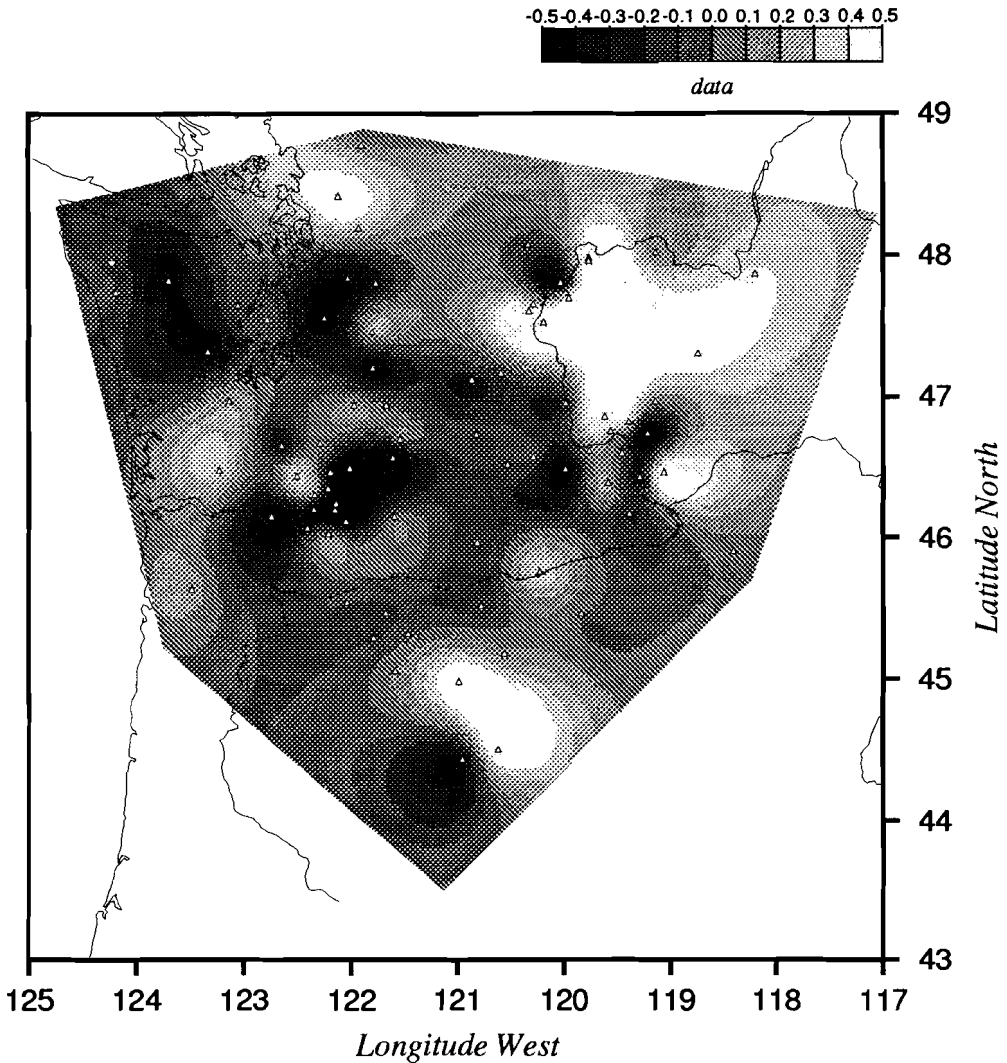


Figure 4. Station statics obtained by averaging observed amplitudes from events with westerly azimuths, plotted at the position of the stations. The station statics are dominated by small-scale variations.

event cluster (of 4 events) with south-easterly back-azimuth the observed amplitude pattern is compared with the predicted amplitude variation across the array. Figure 5 shows the observed pattern, not corrected for station statics. Strong variations are observed, mostly on small scale lengths. This indicates that local effects (including station calibration) dominate the pattern, as the focusing from the 3-D upper mantle structure below the WRSN is not expected to vary over short distances. The corrected amplitudes are shown in figure 6. The dominance of local effects is an important result, as it illustrates that for amplitude studies

## Observed WRSN data from SE (un-calibrated)



*Figure 5.* Observed (uncalibrated) amplitude data for an event with south-easterly azimuth. The focusing due to the subducting slab in this area is expected to be large, since the ray paths for this event travel parallel to the subducting slab. A comparison with figure 3 shows that the amplitudes are dominated by station statics.

station statics must be taken into account, before any inferences about path effects are made. This was already noted by Butler (1984). Averaging over a number of stations, as done by Haddon & Husebye (1978) and Thomson (1983), is not likely to work, since there is no reason to assume that the station statics average out.

The predicted amplitude pattern predicted for the 3-D velocity model of VanDecar (1991) in figure 3 is shown in figure 7. This pattern is obtained using the perturbation

## Observed amplitude pattern from SE; corrected for site statics

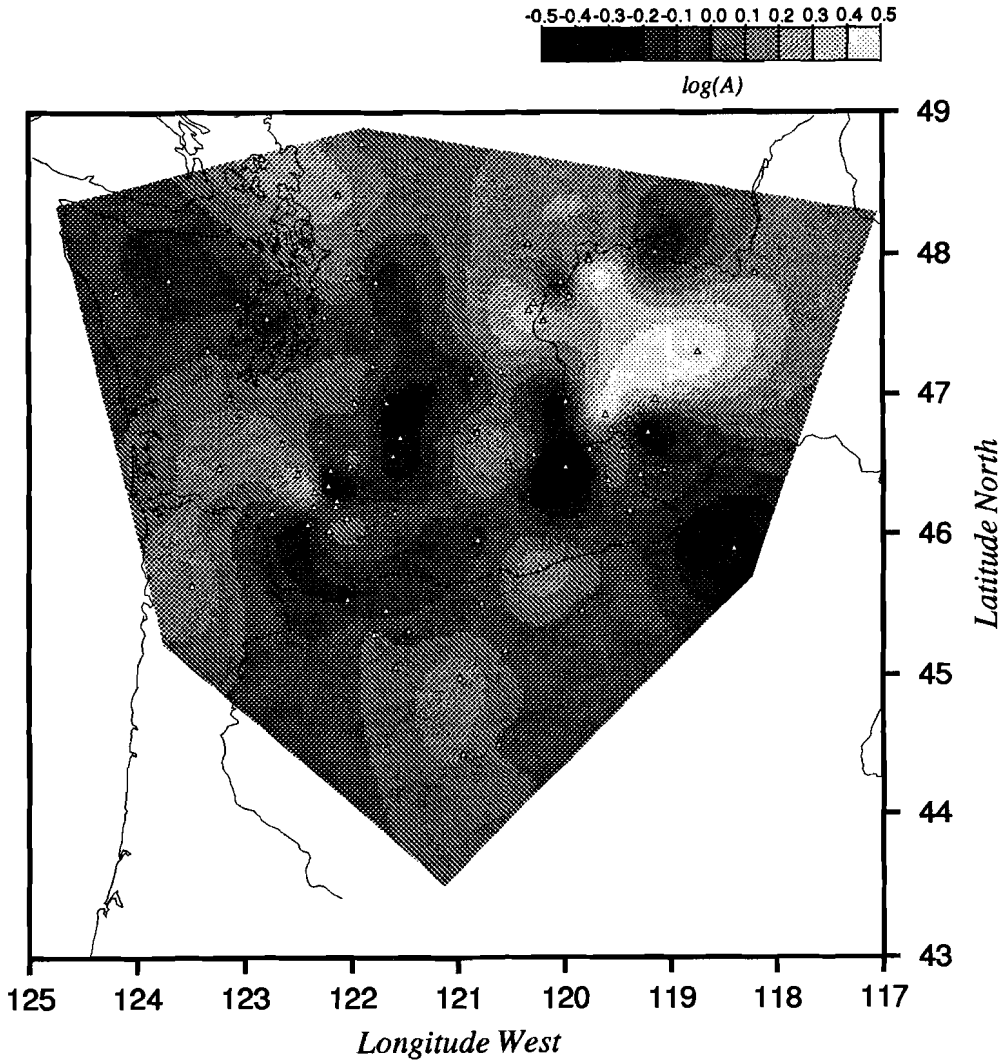


Figure 6. Amplitude pattern from figure 4 corrected for station statics. This pattern is assumed to represent the path effect. The low amplitudes near the centre of the array are attributed to defocusing by the high-velocity slab.

theory described by Neele et al. (1992). Amplitudes are calculated for every event-station pair that is used in the final data set; the predicted amplitudes are then processed exactly as the observed amplitudes. The observed and predicted amplitude patterns seem to correlate on a large scale; differences are present mainly on small scale lengths. The lower amplitudes near the centre of the array can be correlated with the subducting slab: it causes defocusing of rays. The correlation between observed path effects and predicted amplitude pattern from geometrical spreading indicates that focusing and defocusing may dominate the

### Predicted amplitude pattern for UW91F3, event from SE

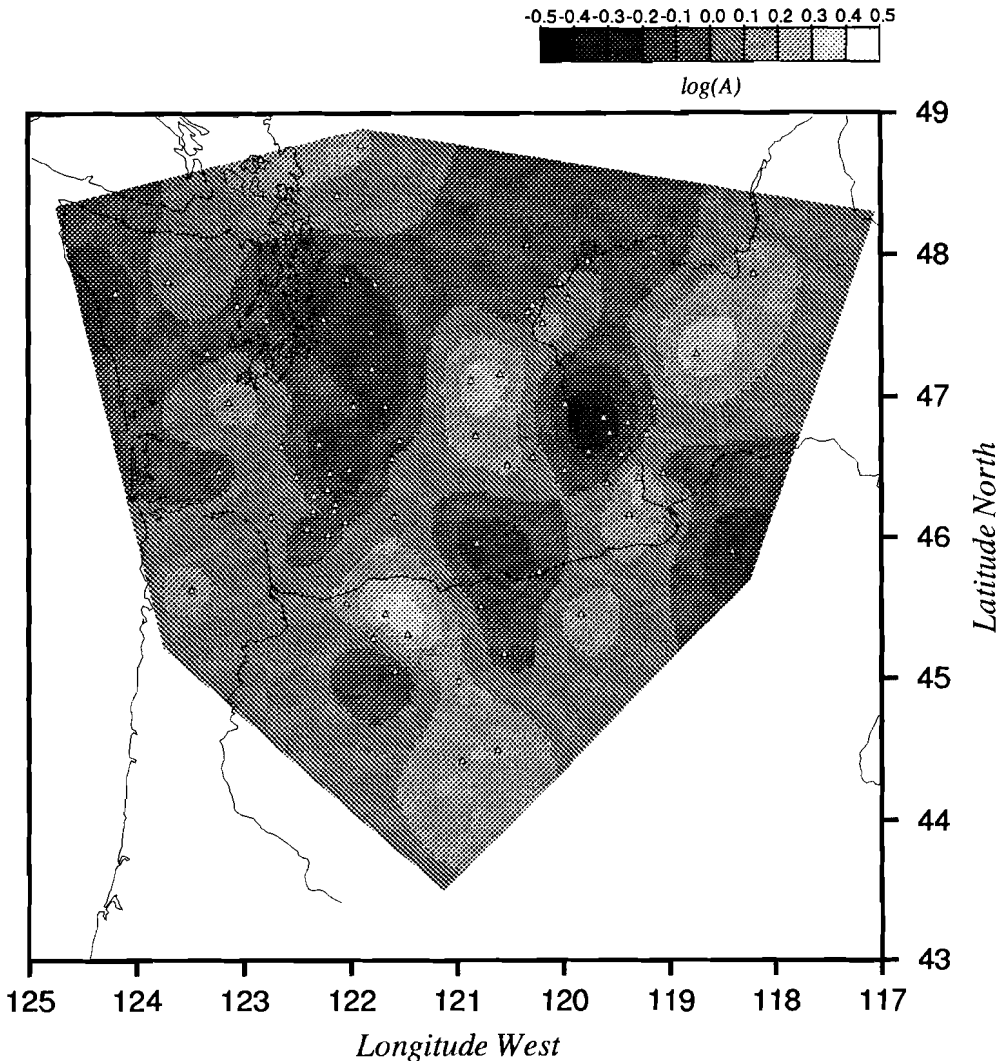


Figure 7. Amplitude pattern for the event from figures 3-5, as calculated from the geometrical spreading in model UW91F3 (figure 9). The low amplitudes near the centre of the array correlate with those in figure 5.

effects of anelasticity in the present data set. This is discussed further in section 4.

### 3. Joint inversion of travel-time and amplitude data

The amplitude data set described in the previous section is now used in a joint inversion with the travel-time data set of VanDecar (1991). It is expected that the amplitude data, being sensitive to the curvature of the slowness field, will help to constrain features such as slab thickness, which can be difficult to resolve using travel times alone. The amplitudes

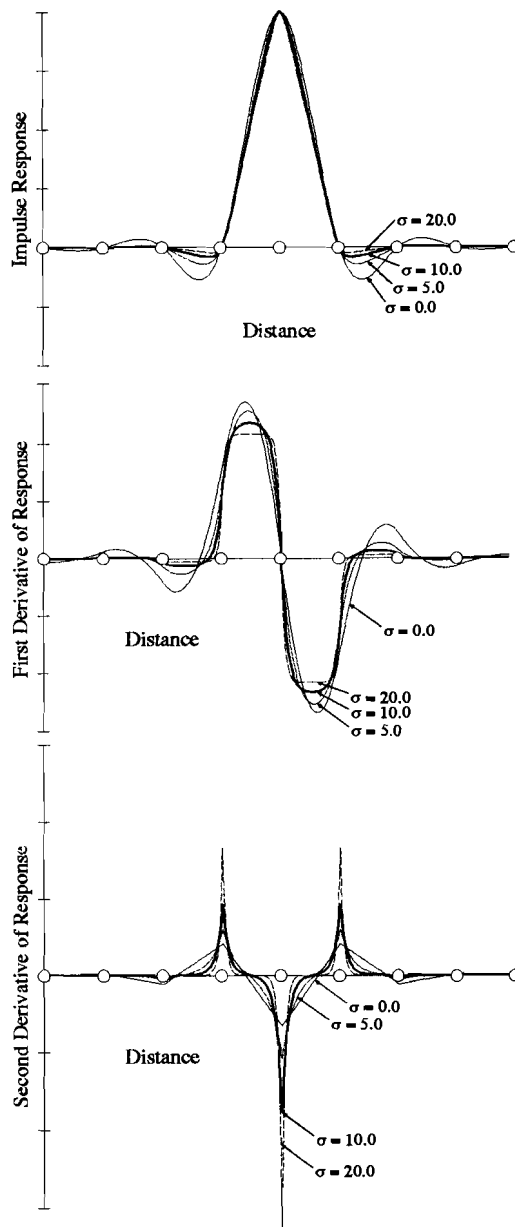
obtained in the previous section are interpreted in terms of geometrical spreading in the upper mantle below the array. The effect of intrinsic attenuation is ignored here. It is shown in the appendix, assuming linear relations between velocity, temperature and  $Q$ , that the effect of attenuation for the present geometry is estimated to be about 10-15% of the observed amplitude fluctuations. As intrinsic attenuation induces a bias to a lower amplitude signal from a high-velocity (high  $Q$ ) subducting slab, ignoring  $Q$  therefore has the effect that in the following inversions the structure necessary to explain the amplitude data may be underestimated by 10-15%.

*Model parameterization.* The region considered here is the upper mantle below WRSN (figure 1). Using short-period  $P$ -wave travel-time data from this array, a tomographic inversion has been performed by VanDecar (1991). In the present study, the inversion method and model parameterization described by VanDecar is used. The upper mantle below the array is parameterized by splines under tension (Cline, 1981), pinned at a series of regular knots. The spline tension controls the interpolation: a zero tension represents cubic splines and a large tension results in tri-linear interpolation. Figure 8 shows a series of interpolation examples with various values for the tension, for a spike of unit magnitude at the central knot. This parameterization yields continuous first and second derivatives, although the second derivative is not smooth. A tension of 10 is chosen here, which minimizes large negative side lobes associated with cubic splines and approaches the localized interpolation of large tension values. This interpolation scheme allows for a smooth slowness distribution and accurate ray tracing through the model. The distribution of knots in the model is shown in figure 9. The model extends in longitude from 128°W to 110°W, in latitude from 42°N to 53°N, and to 900 km in depth. The heavy lines outline the portion of the model in which the highest resolution is expected. In this region the knots are at a distance of 1/3 degree in latitude and 1/2 degree in longitude. In the outer part of the model, set up to absorb any signal from structure outside the region of high resolution, the node spacing is 1° in latitude and 1.5° in longitude. The node spacing with depth is every 25 km to 150 km depth, every 50 km between 150 km and 500 km depth, and every 100 km from 500 km downward. The shortest wave lengths allowed by this parameterization is twice the shortest node separation, or about 50 to 60 km. The number of nodes is 18 in depth, 24 in latitude and 25 in longitude, for a total of 10,800 nodes.

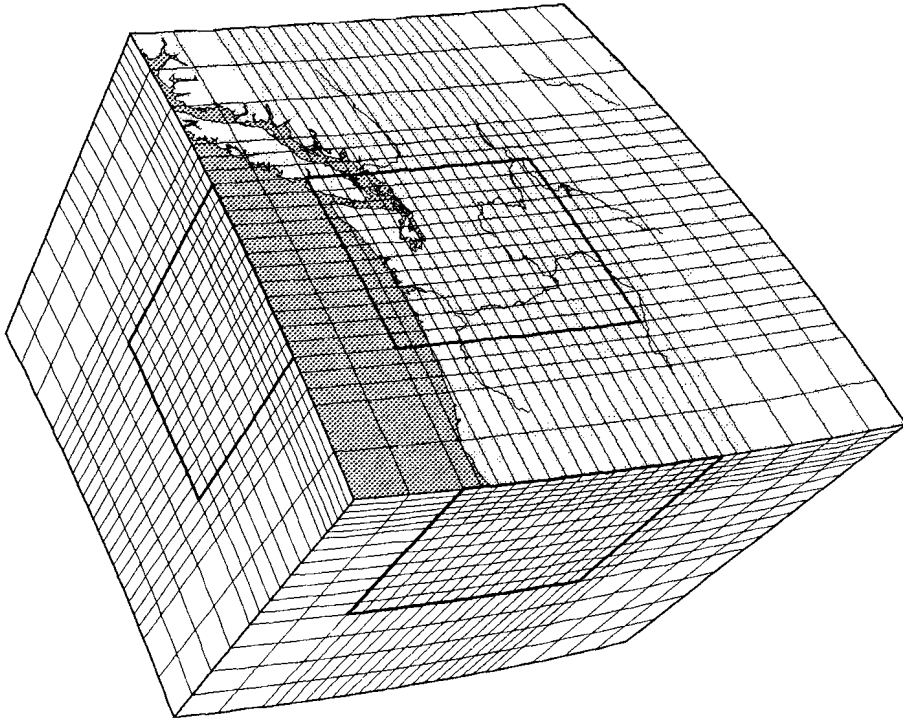
*Background model; travel-time data set.* The background radially-symmetric model is that derived by Herrin (1968), with no first-order discontinuities. The Moho discontinuity in the model is replaced by a smooth transition from 6 km/s at the surface to 8 km/s at 40 km depth. The travel-time data set is that compiled by VanDecar (1991), comprising 8697 teleseismic  $P$ -wave travel times. The noise level is estimated to be about 0.02 s, which follows from the cross-correlation analysis (VanDecar & Crosson, 1990). The signal-to-noise ratio of the travel-time data is therefore approximately 20.

*Inversion method.* The results of an iterative, non-linear travel-time inversion of VanDecar (1991), using the travel-time data set mentioned above, are used here as the starting point. The non-linearity due to ray bending was taken into account by performing a series of linear inversion steps, after each step updating ray paths and travel-time partial derivatives. It was found by VanDecar (1991) that after the third iteration the ray paths did not change





*Figure 8.* Representation of a unit slowness perturbation by splines under tension, for various values for the tension  $\sigma$ . Top: slowness; middle: first derivative; bottom: second derivative.  $\sigma = 0$  represents cubic spline interpolation; a high value results in approximately tri-linear interpolation. The thick line indicates the interpolation for  $\sigma = 10$ , which is used here.

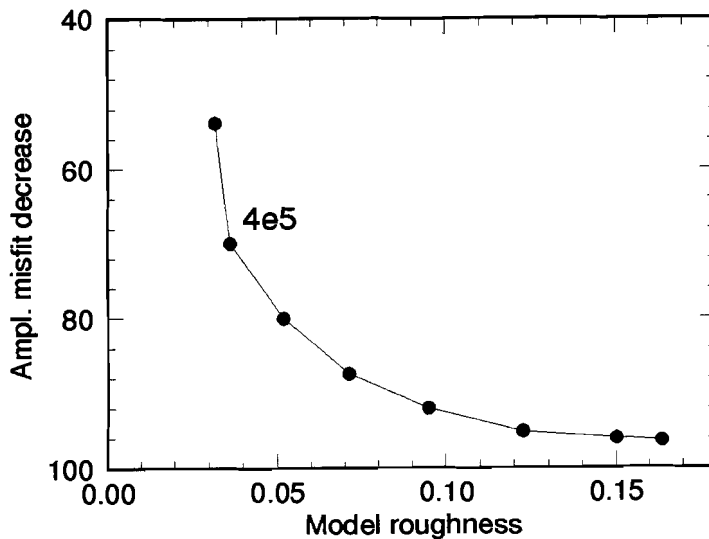


*Figure 9.* Distribution of the knots at which the splines under tension are pinned. There are 18 knots in depth, 25 in longitude and 24 in latitude. The heavy lines represent the projections onto the sides of the portion of the model where the resolution is expected to be highest.

significantly. The model of the third iteration (UW91F3) is shown in figure 3, in a series of vertical slices at constant longitude to a depth of 500 km. The main feature of the model is a high-velocity slab subducting to the east. This feature is interpreted as the subducting Juan-de-Fuca plate. Slowness variations have a scale length of about 100 km and range from -4% to 4%.

In order to minimize the non-linear behaviour of amplitude perturbation as a function of slowness perturbation found by Thomson (1983) and Neele et al. (1992), the model from the third iteration (UW91F3) in the travel-time inversion is used as the starting model for the amplitude data. Ray tracing through model UW91F3 for all event-station combinations of the observed amplitude data set yields the reference ray paths. The ray tracing routine used is basically that of Creager & Jordan (1984), adapted by VanDecar (1991). For each ray path, reference amplitudes and Frechet derivatives of amplitudes for slowness perturbations are calculated using the perturbation theory of Neele et al. (1992); these reference amplitudes and Frechet derivatives are then processed (averaged) in the same way as the observed amplitudes.

The joint inversion is regularized by constraining the second derivative of the model at each node in each of the three directions depth, latitude, longitude independently; in this



*Figure 10.* Trade-off curve of model roughness versus amplitude misfit decrease. The dots represent models obtained with different regularizations and relative weight of the amplitude data. All models produce a travel-time misfit decrease around 74%, which is the same as that produced by UW91F3.

way the smoothest model that best fits the data is found. The regularization constrains the total roughness of the model, rather than minimizing the incremental roughness induced in the joint inversion. A smoothness parameter determines the importance of the model regularization equations relative to the data equations.

In the inversion the weight of the amplitude data relative to the travel times is varied, to find the optimum weight with which the amplitude data can be fit to their variance. In the joint inversion there is one parameter that controls the relative weight of the two data sets, and one to control the importance of the regularization equations. It is chosen here to set these relative weights in such a way, that the total decrease in travel-time misfit is the same for all models obtained in the joint inversions. This is because the travel-time data are of higher quality than the amplitude data. The travel-time data are weighted with the inverse of the estimated variance (0.02 s). In the joint inversion station statics are allowed to absorb any constant amplitude level at a given station (due to, e.g., shallow structure not accounted for by the initial static estimates).

*Results.* Figure 10 shows a trade-off curve of model roughness and amplitude misfit reduction for different levels of regularization and relative weight of the amplitude data. The starting model UW91F3 has an rms roughness of 0.032. All models on the curve explain a travel-time misfit decrease of about 74%. It is apparent that the amplitude data can be fit to an arbitrary level, by lowering the smoothness constraint, and allowing more short-scale structure in the model. Considering the s/n ratio of about 6 in the amplitude data, some models in figure 10 explain a misfit reduction that is not warranted by the variance in the amplitude data. The model for a regularization parameter of  $4 \cdot 10^5$ , which is comparable to

the smoothness constraint used in the travel-time inversions, has almost the same roughness as UW91F3. The differences between this model and UW91F3 are small, have mainly small scale lengths and are located primarily in the deeper parts of the model. The overall slowness variations are not significantly affected. These small-scale differences cause only a slight increase in the roughness of the model, compared to UW91F3, but explain 70% of the amplitude data.

These results indicate that the amplitude data may be fit to an arbitrary degree. The structures necessary to explain the amplitudes appear to be small, short-wavelength fluctuations in the model. At this point it was not clear what causes this result. A possible explanation for these results is a strong sensitivity of the amplitudes on slowness perturbations along the ray path. Small deviations from the starting model may then suffice to explain the data. In the next section joint inversions are performed with synthetic data sets. In this way, the true slowness distribution is known and it can be tested whether the amplitude data increase the resolution of the tomographic inversion. Furthermore, the synthetic amplitude data have much lower variance (only that due to numerical inaccuracies).

#### 4. Joint inversion of synthetic travel-time and amplitude data

Synthetic data are generated for an idealized model of a subducting slab, using the ray paths of the observed data sets. With a known slowness model, it can be tested how the amplitude data, when added to the travel-time data, improve the resolution of the inversion.

*Slab model; synthetic data sets.* The phantom slab model used to generate synthetic data is shown in figure 11. A subducting slab is modeled by a gaussian slowness perturbation extending to about 400 km depth. The slowness perturbation reaches -4%, with a width of 60 km. The background radially symmetric model is again the smoothed Herrin earth model (Herrin, 1968). The event and station distribution for the travel-time data is the same as that of the real travel-time data set used by VanDecar (1991). Synthetic travel-time data for 8697 ray paths are generated by ray tracing through the phantom slab model; random noise with an rms of 0.1 s is added to the data. This is about 5 times the estimated standard deviation that was assigned to the observed travel times. This noise level can be compared with the travel-time signal from the phantom slab, which has an rms of 0.55 s, with a peak at about -3 s. The perturbation theory of Neele et al. (1992) is used to calculate a synthetic amplitude data set, using the event and station distributions from the real amplitude data set. After processing these synthetic data exactly as the observed data, random noise with a rms of 0.01 log(A) is added. Note that this is a much smaller error than present in the observed data. In the joint inversion the data sets are weighted by the inverse rms of the noise, which is the Bayesian approach to the inverse problem. Considering also the conservative estimate of variance in the synthetic travel-time data, the results with these synthetic data sets therefore represent an upper limit on the effect of amplitudes in a joint inversion with travel times.

*Inversion method.* The synthetic data sets are inverted in the same way as the real data sets. Using the travel-time data for the phantom slab model and a radially symmetric model as starting model, four non-linear iterations are performed. A smoothness constraint regularizes the inversion. It is found that after the second iteration the ray paths are close to the

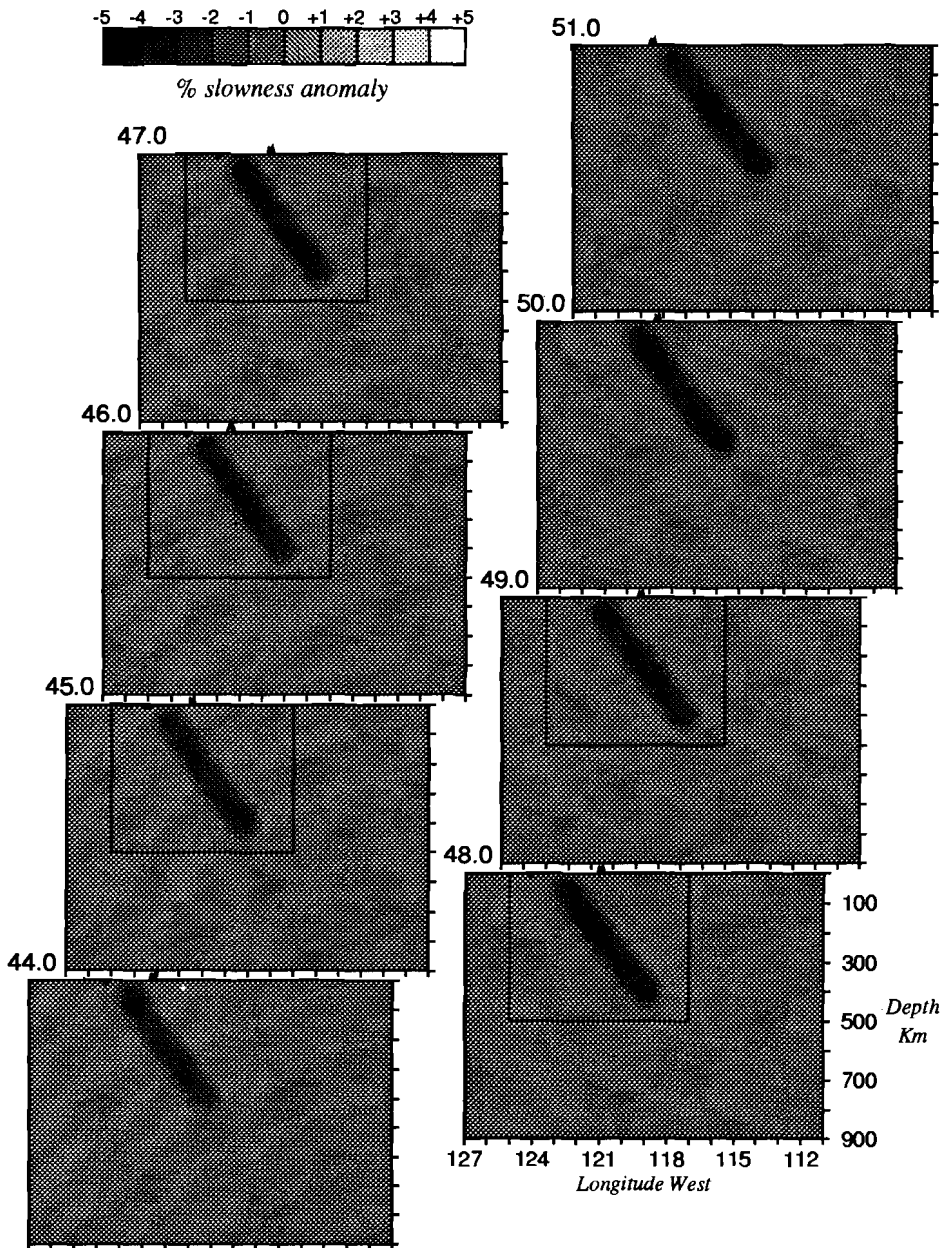


Figure 11. Slowness distribution of the phantom slab model used to generate the synthetic data sets. The slowness scale ranges from -5% to 5%.

true ray paths through the phantom slab model. Figure 12 shows the fourth iteration model, SLAB<sup>(4)</sup>. The travel times have resolved the slab well, although the slab image is a little too wide. The peak slowness of -4% is reached in some places, but especially in the deeper

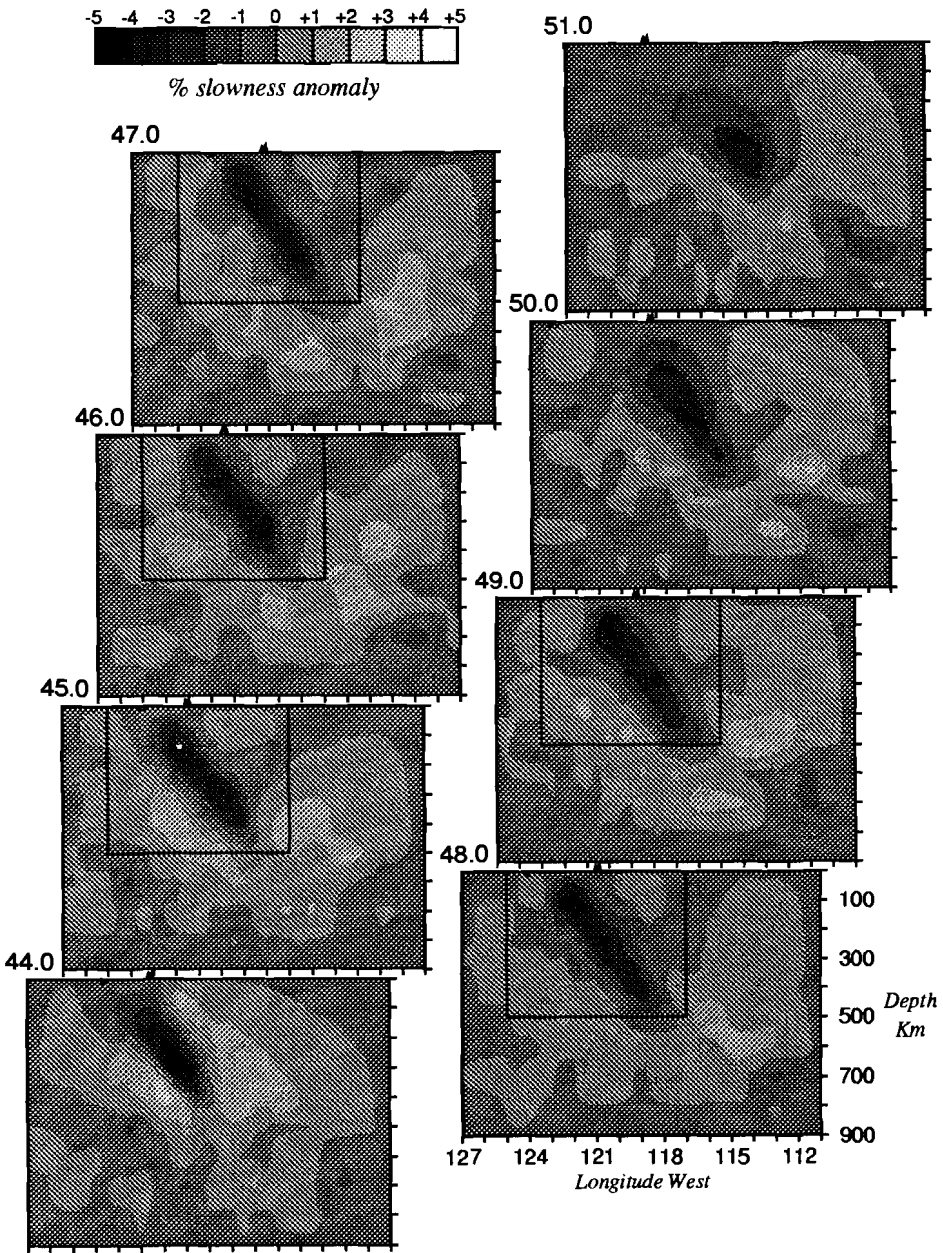


Figure 12. Model SLAB<sup>(4)</sup>, the result of four non-linear iterations of a travel-time tomographic inversion of the data from the phantom slab model in figure 11.

parts of the model the peak slowness is underestimated.

The model from the third iteration, SLAB<sup>(3)</sup>, is used as starting model for a joint inversion of amplitudes and travel times. The non-linearities of the amplitude equations due to

ray shift found by Thomson (1983) and Neele et al. (1992), are expected to be small, since the ray paths in SLAB<sup>(3)</sup> are close to the ray paths in the phantom slab model, for which the synthetic amplitude data were calculated. Frechet derivatives for the amplitude ray paths in SLAB<sup>(3)</sup> are calculated using the perturbation theory of Neele et al. (1992). As in the previous section, in the inversion station statics are allowed to absorb any constant amplitude level at a given station (due to, e.g., shallow structure).

*Results.* Figure 13 shows the result of a joint inversion of amplitudes and travel times. It should be compared with model SLAB<sup>(4)</sup> in figure 12. The models in figures 12 and 13 represent two end-members of a series of models: SLAB<sup>(4)</sup> is obtained with travel times only, whereas figure 13 represents a model obtained with both travel times and amplitudes, with amplitude data of higher quality than the true data set presented in section 2. The model in figure 13 reduces the rms amplitude misfit by 77%, and explains the travel times as well as SLAB<sup>(4)</sup>, at the expense of an increase in roughness of about 1%. The station statics account for about half the misfit reduction; elastic (de)focusing explains the remaining half. An inversion in which no station statics are allowed yielded a model that explains 72% of the amplitude signal, increasing the model roughness only by an additional 3%.

For an event with south-easterly back azimuth the amplitude patterns for the synthetic model and the model from the joint inversion are compared. As the rays from this event travel parallel to the subducting slab, strong (de)focusing is expected. Figure 14a shows the amplitude pattern for the phantom slab model. The low amplitude region in the centre of the array is due to the high-velocity slab, which causes defocusing of rays. The amplitude pattern for model SLAB<sup>(3)</sup>, the starting model in the inversion, is shown in figure 14b. This model predicts an amplitude pattern that has the same long-wavelength structure as the phantom slab model; the differences are caused by the imperfect slab structure in SLAB<sup>(3)</sup>. Figure 14c shows the amplitude pattern for the model from the joint inversion. It is obtained by ray tracing through the model, and represents the true amplitude pattern for the model. It is found that the amplitude pattern found from a linear prediction from the reference model SLAB<sup>(3)</sup> (using the Frechet derivatives for SLAB<sup>(3)</sup>) is almost the same as the true (ray-traced) amplitude pattern shown in figure 14c. This proves that the non-linearities due to ray shift are small. The amplitude pattern in figure 14c shows a good correlation with that in figure 14a.

Figure 15 shows the differences between the phantom slab model, SLAB<sup>(4)</sup>, and the model from the joint inversion in a series of horizontal cross-sections at constant depth and latitude, at depths of 250, 400, 550 and 700 km. The cross-sections in figure 15 represent averages over latitudes from 47° to 48°. It is apparent from the cross-sections that the differences between the models are located mainly in the deeper parts of the model region, but are not larger than 0.5%. The geometrical spreading amplitudes constrain the curvature of the slowness field, and it is apparent from figure 15 that the amplitudes cause small-scale fluctuations in the model. These do not affect the overall slowness variation, but apparently cause a good fit to the amplitude data.

Figure 7 shows that the curvature of the employed splines under tension is localized at the node positions; slight variations in slowness can induce significant variations in the curvature. A parameterization with zero tension (cubic splines) shows a less pronounced and localized effect on the curvature, possibly making it more difficult to change the curvature locally and fit the amplitude data. A joint inversion with a parameterization using cubic

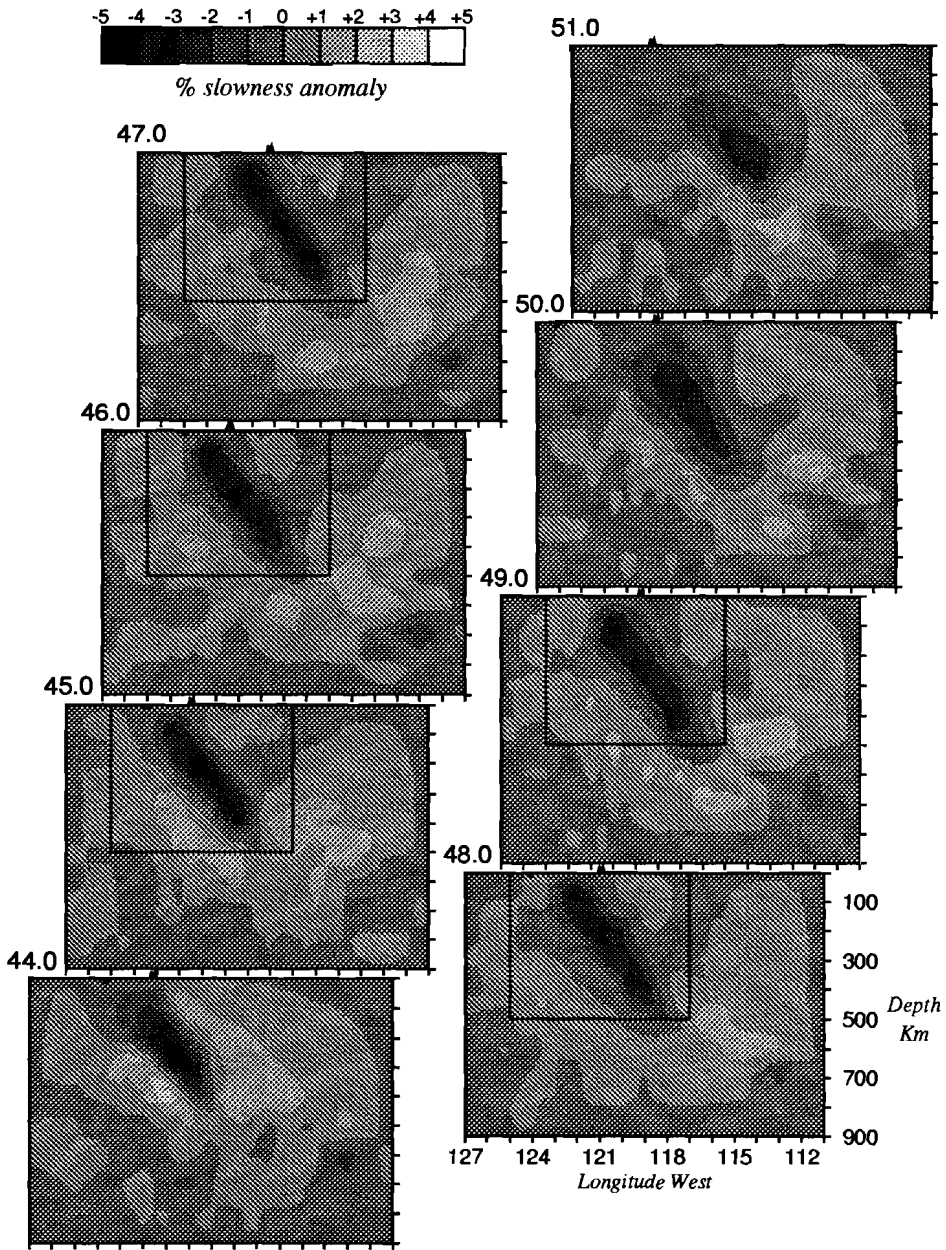


Figure 13. The model obtained from a joint inversion of synthetic travel-time and amplitude data for the phantom slab model, using SLAB<sup>(3)</sup> as a starting model for the amplitude data.

splines, however, does not alter the results: again a decrease in amplitude misfit of almost 70% is found without significantly changing the starting model.

These results imply that small changes to the starting model are sufficient to obtain a



## Amplitude pattern for FSLB, for event from SE

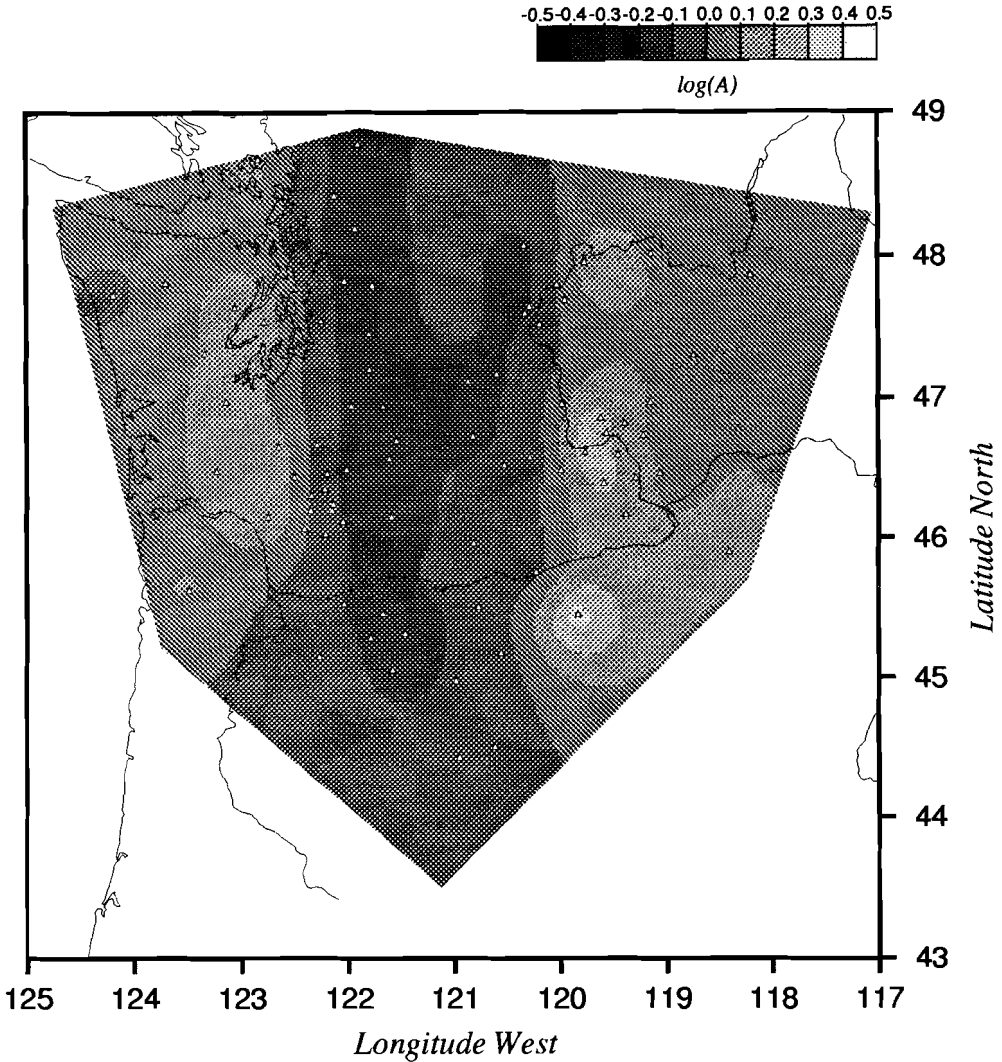


Figure 14a. Synthetic amplitude data for the phantom slab model for an event with south-easterly azimuth. The low-amplitude region near the centre of the array is due to defocusing by the slab.

model that produces the same amplitude patterns as the phantom slab model. This is in agreement with the results with the real data sets. A second experiment dramatically shows the extreme sensitivity of the amplitude data. Starting again with the reference amplitudes and ray paths through model SLAB<sup>(3)</sup>, an inversion is done with the amplitude data set to zero. This means that in this case the model has to be modified in such a way, that it generates no amplitude variation across the array, for all azimuths. The travel-time data are the same as in the first experiment. The model obtained in this joint inversion reduces the rms

## Amplitude data from SLAB(3) (reference amplitudes)

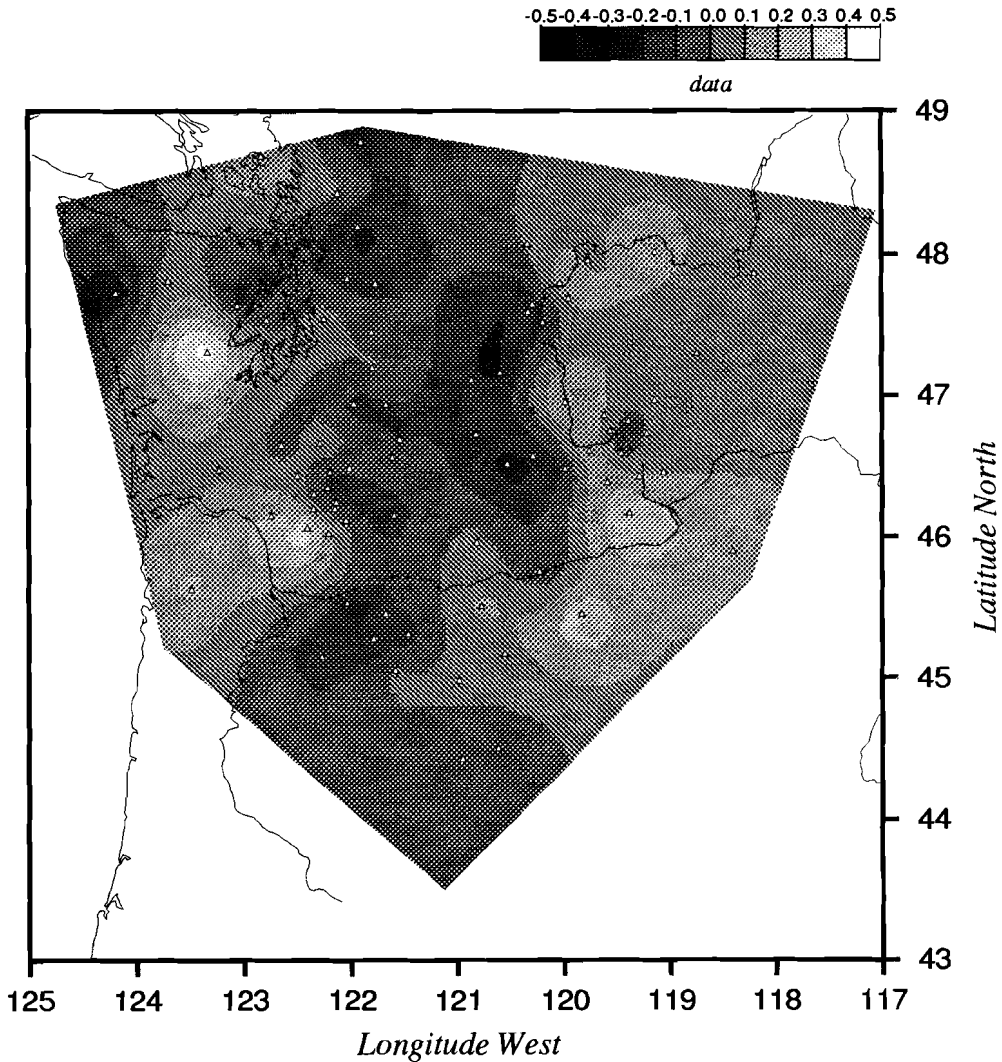


Figure 14b. Synthetic amplitude data for the starting model SLAB<sup>(3)</sup>. The imperfect slab image from the travel-time data is reflected in the distorted region of low amplitudes near the centre of the array.

amplitude misfit by 77%; the travel-time fit is only slightly less than that in the first experiment. The differences between this model and SLAB<sup>(4)</sup> are again small, which is shown in the cross-sections in figure 15. The small-scale slowness fluctuations induced by the amplitude data to remove the amplitude signal from the starting model SLAB<sup>(3)</sup> increase the roughness of the model by only about 1%. The differences between the model from this inversion and the starting model SLAB<sup>(4)</sup> are again located mainly in the deeper parts of the model (see figure 15). This result illustrates the sensitivity of geometrical spreading

## Amplitudes for joint inversion model (along true rays)

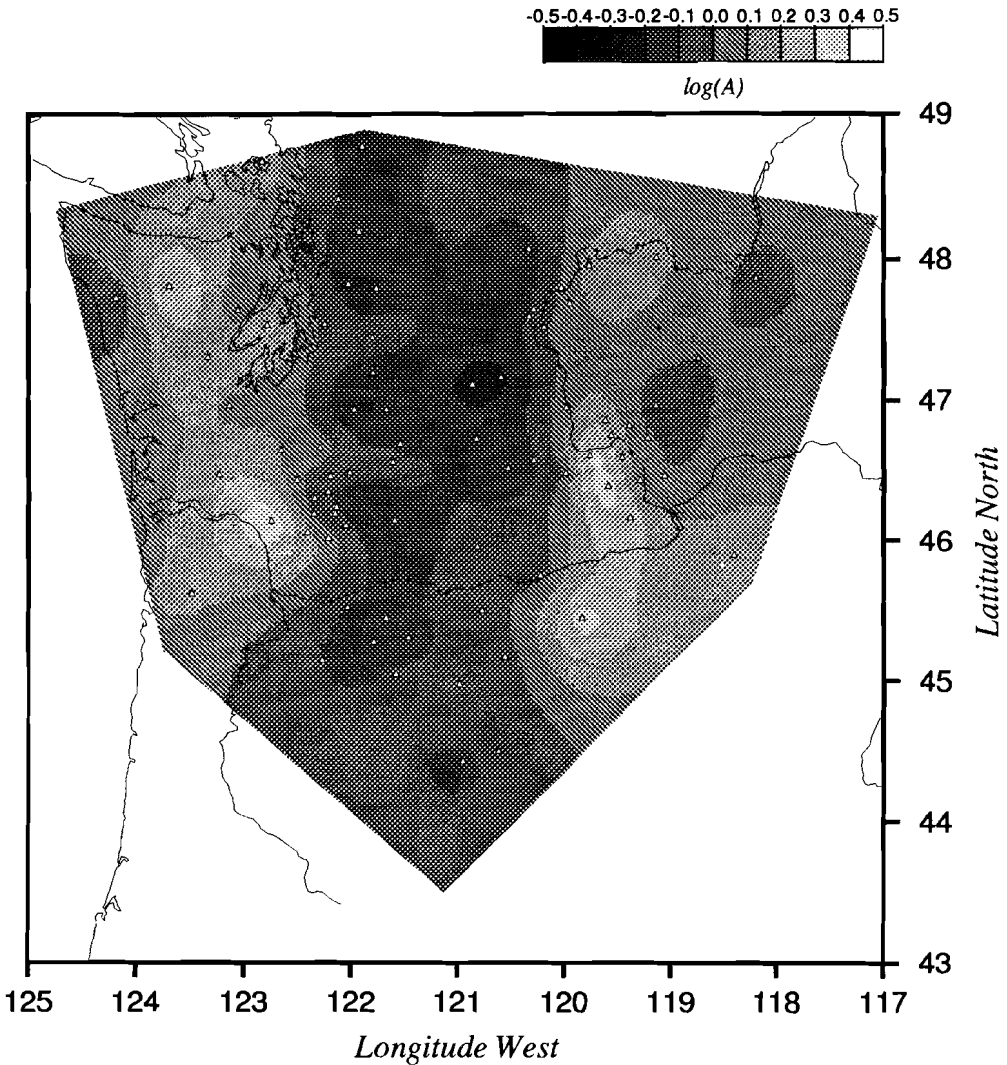


Figure 14c. Amplitude pattern calculated for the model from figure 13.

amplitudes: only small changes to the starting model are necessary to completely reduce its amplitude signature to zero.

These results show that a given amplitude data set can be explained by slightly altering the starting model, which is in agreement with the results obtained with the real data sets. The amplitude data do not resolve the sharp gradients that are difficult to image using teleseismic  $P$ -wave travel times alone, but induce small-scale changes to the starting model, which do not alter the overall slowness variations.

Two factors contribute to this result. Firstly, in this setup the geometrical spreading amplitudes are highly sensitive to slowness perturbations along the ray path. Small

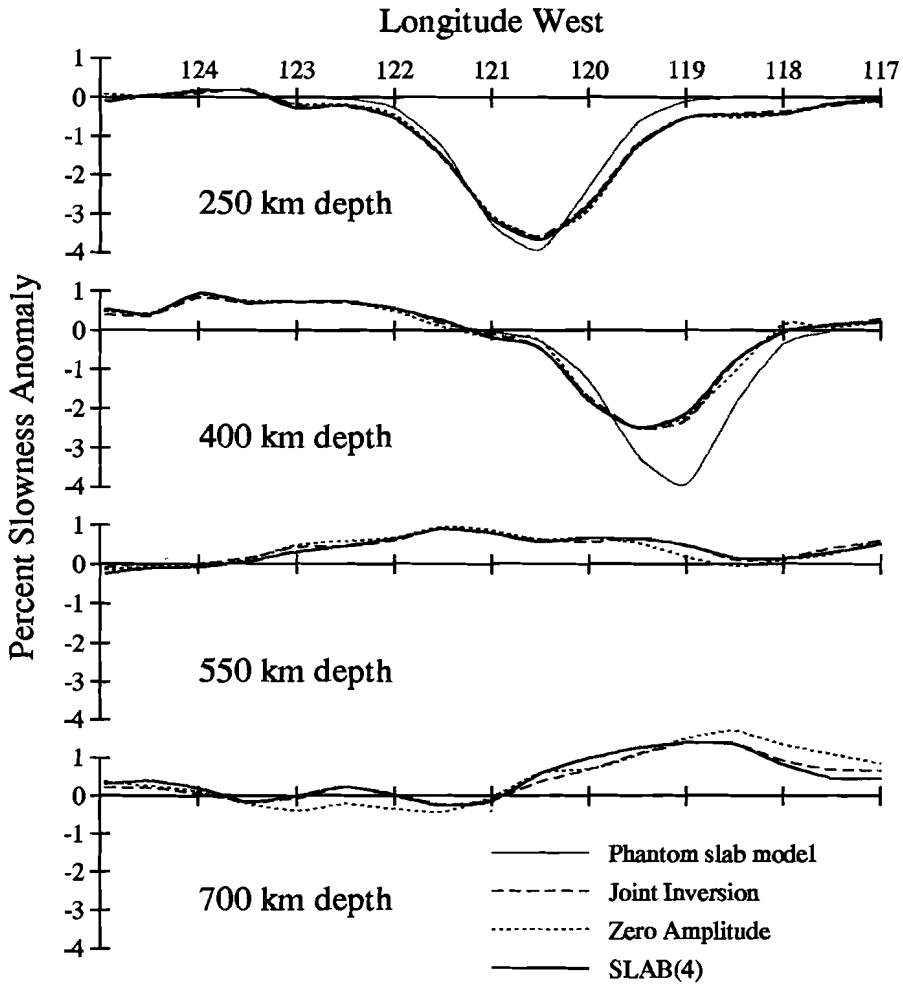


Figure 15. Cross sections at constant depth between longitudes 117°W and 125°W, averaged over latitudes between 47°N and 48°N. Shown are the slowness distributions of the phantom slab model, the fourth-iteration travel-time model SLAB<sup>(4)</sup>, the model from a joint inversion of synthetic data sets for the phantom slab model, and the slab model that is constrained to have zero amplitude signal for all azimuths. The phantom slab model is zero at 700 km depth.

adjustments of the slowness suffice to change the amplitude at the surface significantly. The sensitivity of geometrical spreading amplitudes is controlled by both total ray length and the scale length of the slowness variations that are allowed by the model parameterization. This can be understood from the results of Neele et al. (1992), who show that for a homogeneous reference model the first order amplitude perturbation for a two-point ray can be written as

$$A_1 \approx \int \frac{(S_0 - s)s}{S_0} \partial_{\perp}^2 u_1(s) ds$$

where the integration is along the ray,  $s$  is the distance along the ray from the source to the slowness perturbation,  $S_0$  the total ray length (with  $S_0 - s$  therefore representing the distance along the ray to the station), and  $\partial_{\perp}^2 u_1$  is the second derivative of the slowness perturbation perpendicular to the ray. The sensitivity kernel  $(S_0 - s)s/S_0$  shows that the amplitude perturbation for a slowness perturbation increases with increasing distance from the station, up until the midpoint. This explains why the adjustments to the model are located primarily in the deeper parts of the model. The curvature induced by a unit slowness perturbation depends on the scale lengths allowed by the parameterization, and the geometry of the model setup determines the maximum distance from the station to a perturbed model element. Thus, in the present case of teleseismic rays, a model describing the upper mantle to 900 km depth, and a parameterization allowing slowness variations on a scale of about 60 km, the sensitivity is large.

The second factor is the size of the amplitude data set. The amplitude data set used here is probably too small to induce large-scale changes in the model, since subtler changes affecting individual ray packets suffice to explain the data. The results show that the curvature may be locally changed without altering the overall slowness variations. If a better coverage with rays with amplitude information were available, the data would cause longer-wavelength variations to the starting model and may help increase the resolving power of the tomographic inversion.

The regularization has no influence on the small-scale perturbations to the model, since the increase in model roughness is about 1%. It was shown in the previous section that if the smoothness constraint is relaxed, the amplitude data can be fit to an arbitrary degree.

## 5. Amplitudes as a validation tool.

The results presented in sections 3 and 4 show that the amplitude data set used here is too small to use in a tomographic inversion. This suggests the use of amplitude data as a cross-validation data set. This application of amplitudes was used by, e.g., Cormier (1989) and Gaherty et al. (1991), who used long-period  $S$ -wave amplitudes to validate models of subducting slab structure. This application of amplitudes is tested using both the synthetic and real data sets described above, for the models obtained in different steps of the non-linear inversion of synthetic and real travel-time data, respectively.

For each model tested, the ray perturbation theory of Neele et al. (1992) is used to calculate the amplitudes predicted by the model. The amplitude data sets are processed exactly as the real data set. The cross-correlation coefficient between the predicted and true amplitude data sets (which is either the amplitude data for the phantom slab model or the observed amplitude data) is used as a criterion to measure the validity of a model. The mean is removed from the amplitude response of each event individually. Non-zero means will affect the result of the cross-correlation.

The jackknife method (e.g., Efron, 1982; Lees & Crosson, 1990) is used to obtain an estimate of the variance in the correlation coefficient, which takes the variance in the data into account. With this method, for a set of  $N$  variables,  $N$  estimates of the correlation coefficient are calculated by iteratively leaving out one datum from the set. The  $N$  correlation coefficients  $r_i$  thus obtained are transformed using Fischer's  $z$ -transformation (e.g.,

Snedecor & Cochran, 1967)

$$\mu_i = 1.1513^{10} \log \frac{1+r_i}{1-r_i}$$

The reason for this transformation is that the correlation coefficients have a near normal distribution in the  $\mu$ -domain. From these values  $\mu_i$ , the difference with the  $z$ -transformed correlation coefficient  $\mu_{all}$ , obtained for the total data set, is calculated

$$\hat{\mu}_i = N \mu_{all} - (N - 1) \mu_i$$

The jackknife estimate of the mean ( $\zeta$ -transformed) correlation coefficient is then simply

the average  $\bar{\mu} = \frac{1}{N} \sum_{i=1}^N \hat{\mu}_i$ . The standard error  $\sigma_\mu$  in  $\bar{\mu}$  is

$$\sigma_\mu^2 = \frac{\sum_{i=1}^N \hat{\mu}_i^2 - \frac{1}{N} \left( \sum_{i=1}^N \hat{\mu}_i \right)^2}{N(N-1)}$$

This error is an estimate of the variability of the correlation coefficient due to the variance in the data. The averaged value  $\bar{\mu}$  is transformed back from the  $z$ -domain to obtain the averaged correlation coefficient. In the following, 95% confidence intervals (1.96 times the standard error) around the average correlation value are used.

*Application to synthetic data.* Figure 16a shows the correlation coefficients calculated for the four models SLAB<sup>(i)</sup>,  $i=1-4$ , obtained in successive non-linear travel-time inversion steps using the synthetic travel-time data from the phantom slab model. Starting from the radially symmetric Earth model (plotted at  $x$ -coordinate 0 in figure 16a) with a correlation of almost zero, the correlation coefficient for the consecutive models increases rapidly to about 0.6 for the final model. These correlation coefficients are surprisingly high, given the strong sensitivity of geometrical spreading amplitudes to small changes in the slowness distribution. After the second iteration, the models show a more or less constant correlation, which reflects the result that these models do not differ significantly. The correlation obtained for the model from the joint inversion of travel times and amplitudes is 0.88, reflecting the large decrease in variance explained by this model.

An estimate of the highest possible correlation coefficient that can be obtained with the present amplitude data set, is found by adding a realistic noise level to the synthetic data from the phantom slab model. The correlation coefficient between this noisy data set and the amplitudes from the phantom slab model is 0.53, for an rms noise level of  $0.2 \log(A)$ , which is a reasonable, if perhaps somewhat optimistic estimate of the noise level in real data. This value would be obtained if the travel times would exactly reconstruct the phantom slab model. The correlations between the noisy data set for the phantom slab model and the amplitude data predicted by the slab models SLAB<sup>(i)</sup> are indicated in figure 16a by the squares. The noise level of  $0.2 \log(A)$  decreases the correlation coefficients for the different models by a factor of 2. The difference between the values in figure 16a and the value of 0.53 is due to the imperfect slab reconstruction by the travel times. If noise with an rms value of  $0.3 \log(A)$  is added, the correlation does not exceed 0.25 (indicated by the triangles in figure 16a).

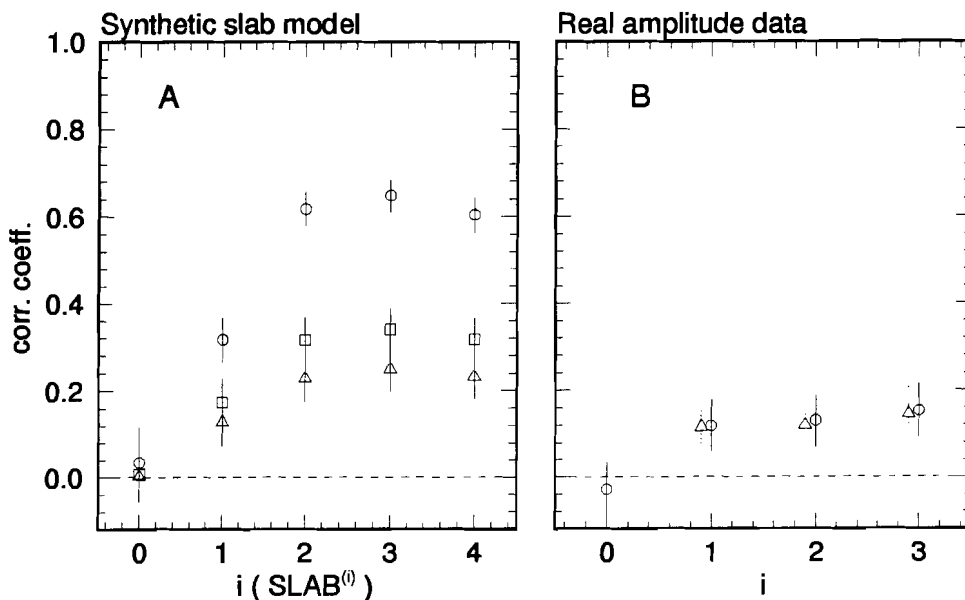


Figure 16. Correlation coefficients for models from successive non-linear iterations of a travel-time tomographic inversion. A) Using the synthetic amplitude data for the phantom slab model; B) Using observed amplitude data. Error bars are calculated with the jackknife method.

*Application to observed data.* The correlation coefficients for the models from the non-linear travel-time inversion of VanDecar (1991), are shown in figure 16b. The models from the three non-linear iterations are indicated by the numbers 1, 2 and 3 on the x-axis; the results for the radially-symmetric starting model are plotted at  $x = 0$ . The results for model UW91F3 (figure 9) are plotted at  $x = 3$ . The circles with solid error bars in figure 16b show the cross-correlation coefficient calculated for the three models. Starting with a small negative correlation for the radially symmetric model, the models from successive iterations produce amplitude patterns that show positive correlation with observed amplitudes. The correlations with the observed data set are significantly above zero (the error bars represent a 95% confidence interval). The correlation coefficients for the models from the different non-linear inversion steps are similar: the observed amplitude data do not prefer either of these models. Nevertheless, there are significant differences between the first iteration model and later models; for example, slab thickness and peak slowness in the slab are significantly altered in the second iteration. The correlation coefficients are comparable to the values obtained when noise is added to the synthetic data (figure 16a). Therefore, a correlation coefficient of 0.16 is near the maximum attainable with the present amplitude data set, given the estimated variance in the data, and the fact that geometrical spreading is the only mechanism affecting amplitudes considered.

A correction for the effect of attenuation to the amplitude patterns predicted for the models from the three non-linear travel-time inversion steps can be made, using the relations between velocity, temperature and  $Q$  presented in the appendix. For each model, corrections for the effect of attenuation are calculated using an oceanic geotherm and the velocity

profile below the northeastern part of each model as reference. As explained in the appendix, this is an upper limit on the effect of  $Q$ . Including  $Q$  has the effect of lowering the amplitudes. The result is shown in figure 16b by the triangles with dotted error bars. The correction for  $Q$  does not change the correlations significantly, which is an indication that the effect of  $Q$  may be small compared to (de)focusing effects.

## 6. Discussion.

In this study the results of a joint tomographic inversion of both real and synthetic  $P$ -wave travel-time and amplitude data for upper mantle slowness distribution are presented. Joint inversions of the real data sets show that the amplitude data can be fit by making small, short-scale slowness changes to the starting model. These fluctuations do not change the overall slowness variations; which ensures that the travel-time fit is unaffected. Synthetic inversions are performed to explain this result. Using an idealized model of a subducting slab, synthetic data are generated, with the same event-station distribution as real data sets. To minimize the (strongly) non-linear behaviour of the amplitude equations for teleseismic distances found by Neele et al. (1992), the synthetic travel times are used in a non-linear tomographic inversion, to obtain a model that is used as the reference model for the amplitude data. The results of a joint inversion show that for this data set the amplitude data do not significantly improve the starting model. The resulting model does not change the travel-time misfit, but explains a reduction of the amplitude misfit by 77%. This is achieved by making small (<0.5%), short-wavelength adjustments to the model, locally changing the curvature of the slowness field, without affecting the larger-scale variations. This is in agreement with the results for the real data sets.

This is due to both the strong sensitivity of the geometrical spreading amplitude to small changes in the velocity distribution and the size of the amplitude data set. Although the strong sensitivity is valuable for the stability of the inversion process, the result is that small changes in the curvature of the slowness field suffice to fit the data, even in cases where the data have no relation to the reference model, as in the second example shown in section 4. With the relatively small number of data used in the amplitude inversion, it appears to be easy to change the curvature of the slowness field on a scale of the node separation, without changing the overall slowness patterns, which have wavelengths of several node spacings. This ensures that the amplitude patterns are explained, without decreasing the travel time fit. This problem should be reduced with a more complete coverage of rays with amplitude data.

The small-scale perturbations induced by the amplitude data increase the model roughness by only a few percent. Therefore, the regularization has no influence on the perturbations. The results for the real data show that the amplitude data can be fit to an arbitrary degree, by sufficiently lowering the regularization. The sensitivity and data size also explain the conclusions of Nowack & Lutter (1988), who showed a dramatic increase in resolving power of tomography if travel-time and amplitude data are combined. These authors used a synthetic model setup with good illumination of the model (leading to a slightly overdetermined problem), and slowness perturbations on a scale of about 1/25 of the total ray length. In the present study this number can be as small as 1/300. In the geometry of Nowack & Lutter (1988) both the sensitivity of amplitudes to slowness perturbations along the ray path is smaller and a more complete coverage of the model is



obtained, resulting in an increase in resolution when the amplitude data are included in the inversion.

It should be noted that the parameterization used in this study does not lead to an overparameterization of the model. The small-wavelength changes to the model induced by the amplitude data are on the scale of the node separation. However, the knot distance is approximately equal to the size of the Fresnel zone of a 1-s *P*-wave. The Fresnel zone should be comparable to the minimum allowable scale length in the model.

The results suggest that with the data sets used, amplitude data do not increase the resolution in the final model and that amplitudes should be used as a validation data set for models obtained by different means. This application is shown for both synthetic and observed data. The travel-time models obtained by VanDecar (1991) predict amplitude patterns that correlate well with the observed amplitude data, given the variance in the data and the fact that only geometrical spreading is considered as a mechanism affecting amplitudes.

The data selection procedure is described in some detail, since the quality of the amplitude data depends strongly on the data processing (removal of outliers, correcting for station statics). As the scatter of *P*-wave amplitudes for closely spaced events observed at a single station is considerable ( $0.3 \log(A)$  is the average variance for the amplitude data in this study, which is a S/N ratio of about 3), in a purely Bayesian inversion these data would be heavily downweighted relative to travel-time data, which generally have much lower variance. A data set 10 times the size of the present amplitude data with the same stations as the present data set (comprising about 40,000 amplitude measurements) would increase the S/N ratio to about 18 through averaging. This is still smaller than the S/N ratio of the travel-time data set of VanDecar (1991) that is used here, which is about 30. Note that apart from increasing the data set to reduce the variance, an even greater increase of the data set would be needed, in order to increase the coverage of the model region with rays with amplitude information.

In conclusion, amplitudes of short-period body waves seem useful in upper mantle velocity inversions only if an extremely large data set is used to both decrease the variance in the data and to obtain a sufficient coverage of rays with amplitude information. However, amplitude data may be useful in, e.g., cross-borehole experiments, where the ratio of ray length and scale length of perturbations is smaller than in the case of upper mantle tomography (Nowack & Lutter, 1988). However, in this case the effect of *Q* may be comparable to that of elastic (de)focusing, making it necessary to take both mechanisms into account. Body wave amplitudes may also be useful in lower mantle tomography. In that case the ratio of ray length and scale length of slowness perturbations may also be much smaller than in the present study, due to the large Fresnel zone of body waves in the lower mantle (Nolet, 1992).

**Acknowledgement.** One of us (J.C.V.) was supported by the U.S. National Science Foundation through Grant no INT-9102113. This research was supported by the Netherlands Organization for Scientific Research (NWO) through the Pionier project PGS 76-144. The University of Washington, and in particular Robert Crosson, Steve Malone and Ruth Ludwig, are gratefully acknowledged for making available the data of the WRSN network.

## References

- Berteussen, K.A., A. Christofferson, E.S. Husebye & A. Dahle, 1975. Wave scattering theory in analysis of *P*-wave anomalies at NORSAR and LASA, *Geophys. J. R. astr. Soc.*, **42**, 403–417.
- Butler, R., 1983. *P*-wave travel time and amplitude in western North America, *Nature*, **306**, 677–678.
- Butler, R., 1984. Azimuth, energy, *Q*, and temperature: variations on *P* wave amplitudes in the United States, *Rev. Geophys. Space Phys.*, **22**, 1–36.
- Butler, R. & L. Ruff, 1980. Teleseismic short-period amplitudes: source and receiver variations, *Bull. Seis. Soc. Am.*, **70**, 831–850.
- Cline, A.K., 1981. *FITPACK - Software package for curve and surface fitting employing splines under tension*, Department of Computer Sciences, University of Texas, Austin, Texas.
- Cormier, V.F., 1989. Slab diffraction of *S*-waves, *J. Geophys. Res.*, **94**, 3006–3024.
- Creager, K.C. & T.H. Jordan, 1984. Slab penetration into the lower mantle, *J. Geophys. Res.*, **89**, 3031–3049.
- Efron, B., 1982. *The jackknife, the bootstrap and other resampling plans*, Soc. for Ind. and Appl. Math. Philadelphia.
- Egbert, G.D. & J.R. Booker, 1986. Robust estimation of geomagnetic transfer functions, *Geophys. J. R. astr. Soc.*, **87**, 173–194.
- Evans, J.R. & J.J. Zucca, 1988. Active high-resolution seismic tomography of compressional wave velocity and attenuation structure at Medicine Lake Volcano, northern California Cascade Range, *J. Geophys. Res.*, **93**, 15,016–15,036.
- Flatte, R.S. & R.S. Wu, 1988. Small-scale structure in the lithosphere and asthenosphere deduced from arrival time and amplitude fluctuations at NORSAR, *J. Geophys. Res.*, **93**, 6601–6614.
- Gaherty, J.B., T. Lay & J.E. Vidale, 1991. Investigation of deep slab structure using long-period *S* waves, *J. Geophys. Res.*, **96**, 16,349–16,367.
- Gudmundsson, O., J.H. Davies & R.W. Clayton, 1990. Stochastic analysis of global traveltimes: mantle heterogeneity and random errors in the ISC data, *Geophys. J. Int.*, **102**, 25–44.
- Haddon, R.A.W. & E.S. Husebye, 1978. Joint interpretation of *P*-wave time and amplitude anomalies in terms of lithospheric heterogeneities, *Geophys. J. R. astr. Soc.*, **55**, 19–43.
- Herrin, E., 1968. 1968 Seismological tables for *P* phases, *Bull. Seis. Soc. Am.*, **58**, 1196–1241.
- Ho-Liu, P., H. Kanamori & R.W. Clayton, 1988. Applications of attenuation tomography to Imperial Valley and Coso-Indian Wells region, southern California, *J. Geophys. Res.*, **93**, 10,501–10,520.
- Huber, P.J., 1981. *Robust Statistics*, John Wiley & Sons, New York.
- Lay, T. & D.V. Helmberger, 1983. Body-wave amplitude and travel-time correlations across north America, *Bull. Seis. Soc. Am.*, **73**, 1063–1076.
- Lees, J.M. & R.S. Crosson, 1989. Tomographic inversion for three-dimensional velocity structure at Mt. St. Helens using earthquake data, *J. Geophys. Res.*, **94**, 5716–5728.
- Moore, B.J., 1980. Seismic ray theory for lithospheric structures with slight lateral variation, *Geophys. J. R. astr. Soc.*, **63**, 671–690.
- Neele, F., J.C. VanDecar & R. Snieder, 1992. A formalism for including amplitude data in tomographic inversions, *Geophys. J. Int.*, submitted.
- Nolet, G., 1992. Imaging the deep Earth: technical possibilities and theoretical constraints, in *Proceedings XXII General Assembly ESC*, edited by A. Roca and D. Mayer-Rosa, pp. 107–114.
- Nowack, R.L. & W.J. Lutter, 1988. Linearized rays, amplitude and inversion, *Pure Appl. Geophys.*, **128**, 401–421.
- Ringdal, F., 1977. *P*-wave amplitudes and sources of scattering in Mb-estimates, *J. Geophys.*, **43**, 611–622.
- Ringdal, F., E.S. Husebye & A. Dahle, 1972. Event detection problems using a partially coherent seismic array, *NTNF/NORSAR Rep.* 45.
- Sato, H., I.S. Sacks, T. Murase, G. Muncill & H. Fukuyama, 1989. *Q<sub>p</sub>*-melting temperature relation

- in peridotite at high pressure and temperature: attenuation mechanism and implications for the mechanical properties of the upper mantle, *J. Geophys. Res.*, **94**, 10,647–10,661.
- Snedecor, G.W. & W.G. Cochran, 1967. *Statistical methods*, Iowa state press, Ames, Iowa.
- Spakman, W., 1991. Delay time tomography of the upper mantle below Europe, the Mediterranean, and Asia Minor, *Geoph. J. Int.*, **107**, 309–332.
- Spakman, W., S. Stein, R. van der Hilst & R. Wortel, 1989. Resolution experiments for NW Pacific subduction zone tomography, *Geophys. Res. Lett.*, **16**, 1097–1100.
- Takahashi, E., 1986. Melting of a dry peridotite KLB-1 up to 14 GPa: implications on the origin of peridotitic upper mantle, *J. Geophys. Res.*, **91**, 9367–9382.
- Takahashi, E., 1990. Speculations on the Archean mantle: missing link between komatiite and depleted garnet peridotite, *J. Geophys. Res.*, **95**, 15,941–15,954.
- Thomson, C., 1983. Ray-theoretical amplitude inversion for laterally varying velocity structure below NORSAR, *Geophys. J. R. astr. Soc.*, **74**, 525–558.
- Thomson, C.J. & D. Gubbins, 1982. Three-dimensional lithospheric modelling at NORSAR: linearity of the method and amplitude variations from the anomalies, *Geophys. J. R. astr. Soc.*, **71**, 1–36.
- VanDecar, J.C., 1991. *Upper mantle structure of the Cascadia subduction zone from non-linear teleseismic travel-time inversion*, PhD thesis, University of Washington, Seattle, Washington.
- VanDecar, J.C. & R.S. Crosson, 1990. Determination of teleseismic relative phase arrival times using multi-channel cross correlation and least squares, *Bull. Seis. Soc. Am.*, **80**, 1049–1061.
- Van der Hilst, R.D., and E.R. Engdahl, 1991. On ISC PP and pP data and their use in delay-time tomography of the Caribbean region, *Geophys. J. Int.*, **106**, 169–188.
- Zhou, H.W. & R.W. Clayton, 1990. *P* and *S* wave travel time inversions for the subducting under the island arcs of the northwest Pacific, *J. Geophys. Res.*, **95**, 6829–6852.

### Appendix. The effect of attenuation.

Ignoring the effect of anelastic damping may not be warranted in a subduction zone region, where partial melt is expected in the mantle above the subducting slab. Furthermore, the far western part of the region below the WRSN is oceanic, whereas continental crust and upper mantle is found in the eastern parts of the region. Different temperature regimes throughout the upper mantle below the WRSN region may have an appreciable effect on the *P*-wave amplitudes.

An upper limit on the effect of *Q* can be obtained by interpreting the slowness variations present in the model UW91F3 as the result of temperature variations, and converting these to a *Q* structure. Once the *Q*-distribution is computed, the attenuation along a ray *i* is given by the line integral (see, e.g., Ho-Liu et al., 1988)

$${}^{10}\log\left(\frac{\tilde{A}}{A}\right) = -\frac{1}{\epsilon \log(10)} \int \pi f \frac{dl}{Q(r_i)\alpha(r_i)} \quad (4)$$

where  $\tilde{A}$  is the perturbed amplitude and *A* the original (reference) amplitude; *f* is the frequency of the wave;  $\alpha$  is the *P*-wave velocity.

An upper limit for the effect of *Q* can be found by calculating (4) for a ray through that part of the model where the lowest *Q* values are found. This is the case for the north-western part of the model, where young oceanic lithosphere reaches the subduction zone. The pronounced low velocities in this region are interpreted as a low-velocity layer that is present under young oceanic regions, but not in the reference radial velocity model (VanDecar, 1991). It is therefore not reasonable to interpret the low velocities as anomalous, as they are an artifact of the radial reference model and would yield unrealistically low *Q* values. For the calculation of *Q* in this region an oceanic geotherm for an age of 20 Ma and a

potential mantle temperature of 1400 K is used; the total velocity profile beneath this region (i.e., radial velocity plus perturbations), including a low-velocity zone at about 100 km depth, is used as reference. The  $Q$ -variation with depth is calculated using the results of Sato et al (1989). These authors obtained an empirical relation between the homologous temperature  $T_m/T$ , with  $T_m$  the melting temperature, from high-pressure laboratory experiments on peridotite. The melt curve used is that given by Takahashi (1986, 1990).  $Q$ -values thus obtained reach as low as 100 in the low-velocity zone. Below about 300-400 km  $Q$  becomes so large that the wave becomes effectively insensitive to  $Q$  variations. The maximum attenuation for a ray through the oceanic lithosphere is about -0.13 log relative amplitude (the minus sign indicating a lower amplitude). Rays that traverse the slab and the continental lithosphere are less affected by  $Q$  since these regions are cooler and thus have higher  $Q$  values. The lower limit of -0.13 (the upper limit is obviously zero) must be compared to the observed peak-to-peak log amplitude variations (corrected for site response) of about 1.0. Therefore, using realistic values for velocity and temperature the maximum effect of  $Q$  structure is at most 10-15% of the amplitude variations observed.

---

## Samenvatting

De laatste jaren stijgt het aantal seismische stations dat registraties van goede (digitale) kwaliteit produceert. Zowel kort- als langperiodische data worden op routinematige wijze aan centrale databanken gestuurd. Ondanks de stijgende hoeveelheid stations wordt de globale seismologie, die de structuur van de aarde tracht te ontrafelen, nog steeds gehinderd door een onevenwichtige spreiding van stations en bevingen. Het gevolg is dat er slechts een beperkte kennis van het seismische golfveld bestaat, hetgeen het bepalen van de structuur van de aarde bemoeilijkt. Het is daarom noodzakelijk om zoveel mogelijk informatie uit een seismogram te halen.

Van een seismogram wordt echter maar een klein gedeelte op min of meer routinematige wijze gebruikt. De meest gebruikte delen van het seismogram zijn de reistijd van de *P*-golf en de golfrein van de oppervlaktegolven. De *P*- en *S*-reistijd wordt in grootschalige inversies gebruikt om modellen van snelheidsvariaties in de aarde te maken. De (langperiodische) oppervlaktegolven worden gebruikt om modellen van de bovenmantel te construeren met behulp van mode-sommatie. Recente ontwikkelingen laten zien dat het mogelijk is om ook de *S*-golf in deze golfvormaanpassing op te nemen. Dit blijkt vooralsnog niet mogelijk met *P*-golven; de oorzaak is een te lange rekentijd om een *P*-golf met een som van mantel-modes te beschrijven.

Grote delen van het seismogram blijven dus nog ongebruikt in inversies voor aardstructuur, zoals de bovengenoemde golfvormaanpassing. Deze delen vormen in principe een rijke bron aan informatie. Indien toegevoegd aan al gebruikte gegevens, kan deze extra informatie de huidige kennis van de structuur van de aarde verbeteren. Het onderzoek gepresenteerd in dit proefschrift heeft als doel om die delen van het seismogram, die niet vaak gebruikt worden, te bestuderen, om te bepalen wat het informatiegehalte is en welke klasse van modellen nodig zijn om deze data te kunnen beschrijven (bijvoorbeeld een gelaagd tegenover een lateraal heterogeen model). Deze kennis is nodig voordat eventuele inversiemethoden voor deze data kunnen worden opgezet. Indien mogelijk worden methodes ontwikkeld om de informatie uit de data te halen.

In hoofdstukken 2 en 3 worden de zogenaamde coda intervallen tussen de ruimtegolven (*P*, *S*, etc.), die een aanzienlijk gedeelte van het seismogram beslaan, bestudeerd. De golfvormen van de ruimtegolven zelf worden onder andere gebruikt om modellen van de gelaagdheid van de bovenmantel te construeren. Uit vele studies blijkt dat golfverstrooiing de coda-intervallen van kortperiodische seismogrammen domineert. Dit heeft tot gevolg dat de relatie tussen data en aardstructuur zeer complex kan zijn. De aandacht wordt

daarom in hoofdstukken 2 en 3 op langperiodische data gericht, omdat van deze data verwacht wordt dat verstrooiing minder belangrijk is. Een analyse van data van twee seismische arrays in Europa (het Netwerk van Autonom Registerende Seismografen (NARS) van Zweden tot Spanje en het netwerk bij Graefenberg in Duitsland) laat zien dat het karakter van de coda na de  $P$ -golf duidelijk verschilt van dat van de coda van  $PP$ - en  $S$ -golven. De  $P$ -coda vertoont een coherent gedrag tussen de stations van het relatief kleine GRF, terwijl de coda van de andere ruimtegolven gedomineerd wordt door incoherente energie. Dit laatste duidt op golfverstrooiing. Met behulp van een nieuwe beamforming-methode, speciaal toepasbaar op de relatief korte tijdsintervallen in ruimtegolfddata, kan het signaal van een array worden geanalyseerd. Dit geeft informatie over zowel het type golven in de coda als de richting waaruit de energie aankomt. Het blijkt dat de verstrooide energie in de coda van  $PP$ - en  $S$ -golven kan worden toegeschreven aan oppervlaktegolven, die uit verschillende richtingen binnenkomen. Om zo vroeg in een seismogram te arriveren, moeten deze oppervlaktegolven opgewekt zijn door de ruimtegolven. Berekeningen laten zien dat het mogelijk is het energieniveau in de coda-intervallen te verklaren met verstrooiing en conversie van ruimtegolven naar oppervlaktegolven aan relief aan het oppervlak. De  $P$ -coda bevat weinig verstrooide energie; synthetische seismogrammen laten zien dat het karakter van deze coda voor kleine arrays (Graefenberg) kan worden verklaard met gelaagde, lateraal homogene modellen van de bovenmantel. Het incoherente gedrag van de  $P$ -coda op grotere schaal (NARS) kunnen dan verklaard worden met variaties in deze gelaagdheid. De synthetische seismogrammen suggereren het gebruik van de  $P$ -coda om bijvoorbeeld variaties in de lage-snelheidslaag te bestuderen. Bovendien is het misschien mogelijk om met deze coda vast te stellen of er een lage-snelheidslaag in de  $P$ -snelheid bestaat.

De resultaten van hoofdstukken 2 en 3 geven aan dat de coda van  $PP$ - en  $S$ -golven beschreven moeten worden met golfvoortplanting in een lateraal heterogene aarde, waarin golfverstrooiing en golfconversie een complexe coda veroorzaken. De informatie in deze delen van het seismogram (ondiepe structuur) is waarschijnlijk op eenvoudiger wijze uit andere data te halen. De coda van  $P$ , daarentegen, kan beschreven worden met golfvoortplanting in een gelaagd aardmodel en lijkt bruikbaar om de bovenmantel tussen beving en station te karakteriseren.

In hoofdstuk 4 wordt een reflectie aan de 400-km discontinuïteit ( $P_{400}P$ ) in de  $P$ -coda gebruikt om deze discontinuïteit in de bovenmantel onder Hawaï te bestuderen. De variaties in amplitude van  $P_{400}P$  voor verschillende beving-station combinaties worden geïnterpreteerd als het resultaat van focusering of defocusering door variaties in diepte van de discontinuïteit. Door gebrek aan gegevens is het niet mogelijk om een deterministisch model te construeren, maar met behulp van synthetische seismogrammen kan wel een schatting gemaakt worden van de combinaties van golflengtes en amplitudes van de dieptevariaties van de discontinuïteit die in overeenstemming met de data zijn. Dieptevariaties met golflengtes in het bereik van 500-1500 km en amplitudes van ongeveer 15-20 km kunnen waargenomen amplitudevariaties van  $P_{400}P$  verklaren. Als meer data voorhanden zijn, zijn  $P_{400}P$  golven wellicht bruikbaar om laterale variaties van de 400-km discontinuïteit in kaart te brengen. Deze kunnen bijvoorbeeld veroorzaakt worden door bovenmantelconvectie. Een interessant resultaat van de berekeningen is dat de amplitude van  $P_{400}P$  als functie van de positie van het straalgeometrische reflectiepunt kan variëren op een schaal die kleiner is dan de grootte van de Fresnelzone. De oorzaak hiervoor is, dat,

door de geringe kromming van het reistijdoppervlak van  $P_{400}$   $P$  bij het reflectiepunt, relatief gering relief van de discontinuïteit meerdere punten van stationaire reistijd genereert. Deze geven aanleiding tot meerdere, met elkaar interfererende aankomsten in het seismogram, met mogelijk sterk variabele amplitude tot gevolg.

In de laatste twee hoofdstukken van dit proefschrift wordt aandacht besteed aan de amplitude van kortperiodische  $P$ -golven. Hoewel de amplitude van een  $P$ -golf relatief eenvoudig te meten is, wordt deze grootte weinig gebruikt. De oorzaken hiervoor zijn de grote variabiliteit van de amplitude (kleine variaties in bron-ontvanger combinaties kunnen grote variaties in de amplitude veroorzaken) en de complexe relatie tussen amplitude en snelheidsstructuur van de aarde. De  $P$ -golf amplitude draagt echter in principe waardevolle informatie over de aardstructuur, die onafhankelijk is van, bijvoorbeeld, reistijdinformatie. Een tomografische inversie van reistijd- en amplitude-data kan het oplossend vermogen van de inversie vergroten. Met een tomografische inversie zijn in het algemeen grote datasets gemoeid, en het is uit eerdere studies al gebleken dat vooral de amplitudes veel rekentijd vergen.

In hoofdstuk 5 wordt een formalisme gepresenteerd waarmee op efficiënte wijze Frechet afgeleiden van amplitude naar snelheidsverstoringen langs de straal kunnen worden berekend. Hiermee is het mogelijk om grote hoeveelheden amplitude-data te verwerken. De Frechet afgeleiden geven de gevoeligheid van een amplitude voor snelheidsvariaties langs het pad van de  $P$ -golf. Berekeningen met een realistisch model van de snelheidsverdeling in de bovenmantel laten het sterk niet-lineaire gedrag zien van amplitudevariaties als functie van snelheidsperturbaties. Hierdoor is de gelineariseerde theorie van hoofdstuk 5 toepasbaar voor slechts een beperkt gebied van snelheidsperturbaties ten opzichte van een gegeven referentiemodel.

De ervaring uit hoofdstuk 5 wordt in hoofdstuk 6 gebruikt in een inversie van waargenomen  $P$ -golf amplitudes naar snelheidsvariaties in de bovenmantel, samen met  $P$ -golf reistijden. Het gebied waarop de methode toegepast wordt is de bovenmantel onder de staat Washington in de Verenigde Staten, waar de data geregistreerd zijn. De verwerking van de amplitudes is belangrijk in het verkrijgen van een betrouwbare dataset. Een grote hoeveelheid  $P$ -golf amplitudes is verzameld, zodat door middeling van de data de aanzienlijke variabiliteit in de amplitudes verminderd kan worden en de betrouwbaarheid vergroot. Dit gebeurt met een methode, die zeer afwijkende metingen (zogenaamde 'outliers') identificeert en verwijdert. Verder blijkt uit de dataverwerking dat de amplitude vooral beïnvloed wordt door effecten vlakbij het station. Dit zijn bijvoorbeeld de versterkingsfactor van het instrument, locale structuur onder het station. Deze effecten blijken het effect van golfvoortplanting door de mantel te domineren.

Als startmodel in de inversies wordt een model genomen dat uit de reistijden is verkregen. Op deze manier wordt de sterke niet-lineariteit, die in hoofdstuk 5 werd gevonden, geminimaliseerd. Het resultaat van de inversie is dat met kleine aanpassingen aan het model de amplitude-data verklaard kunnen worden. Inversies met synthetische data bevestigen deze uitkomst. Deze resultaten lijken erop te duiden dat  $P$ -golf amplitudes niet helpen bij het oplossen van de snelheidsstructuur van de bovenmantel. Een aantal factoren speelt een rol in de verklaring van dit resultaat. De eerste factor is de grote gevoeligheid van de amplitudes voor snelheidsverstoringen. Deze wordt veroorzaakt door de combinatie van de grote lengte van het looppad van een teleseismische  $P$ -golf met de kleine schaal van snelheidsvariaties die toegelaten worden door het model van de bovenmantel. Hierdoor zijn

kleine verstoringen van het startmodel voldoende om de amplitude-data te verklaren. Ten tweede is de amplitude-dataset niet groot genoeg om een goede stralenbedekking van het model te krijgen. De amplitudes leggen de tweede afgeleide van de snelheid vast langs het golfpad; het blijkt gemakkelijk om de tweede afgeleide plaatselijk in het model te veranderen, zonder de grootschalige snelheidsvariaties (vastgelegd door de reistijden) aan te tasten. Een betere stralenbedekking is nodig om met de amplitudedata coherente veranderingen in het model te veroorzaken. Hierbij moet vermeld worden dat de variantie in amplitudes aanzienlijk groter is dan de variantie in reistijden. Een zeer grote amplitude-dataset is nodig om de variantie zodanig te reduceren dat de signaal-ruis verhouding van de amplitudes vergelijkbaar is met die van de reistijden. In een bayesiaanse inversie worden de data geschaald met de inverse variantie en zouden de twee datasets dan gelijk gewicht toegekend worden. De dataset uit hoofdstuk 6 zou een factor tien groter moeten zijn (ongeveer 50.000 amplitudedata) om deze situatie te benaderen; een nog grotere dataset is nodig om de stralenbedekking van het model te verbeteren.

De amplitude van kortperiodische  $P$ -golven lijkt dus bruikbaar in een tomografische inversie voor bovenmantelstructuur, mits de dataset zodanig groot is, dat zowel een goede bedekking van het model als een grote variantiereductie behaald wordt. De  $P$ -golf amplitude heeft waarschijnlijk meer waarde in experimenten waarbij de gevoeligheid van de amplitudes voor kleine verstoringen in het snelheidsmodel kleiner is. Hierbij kan gedacht worden aan 'cross-borehole' data, waarbij de verhouding tussen straallengte en schaallengte van de snelheidsvariaties aanzienlijk kleiner is dan in bovenmanteltomografie met kortperiodische golven. Verder kan de amplitude waardevolle informatie verschaffen in studies met langperiodische oppervlaktegolven, waarvoor hetzelfde kan gelden.



## Acknowledgements

Het onderzoek zoals beschreven in dit proefschrift is tot stand gekomen met de hulp en ondersteuning van velen, die ik op deze plaats wil noemen.

Allereerst wil ik mijn begeleider, Roel Snieder, bedanken. Zijn enthousiasme en interesse in mijn werk waren steeds een bron van inspiratie. Zijn diepe inzicht en intuïtie in theoretische en praktische problemen gaven altijd een aanzet in de goede richting.

I am greatly indebted to John VanDecar, whose cooperation was crucial for the research presented in chapters 5 and 6. He generously shared his data-analysis and inversion programs and was always ready to listen. It was a pleasure to work with him.

I gratefully acknowledge the University of Washington, Seattle, and in particular Robert Crosson, Steve Malone and Ruth Ludwig, for making available the data of the WRSN network that were used in chapters 5 and 6.

Ik heb de vele discussies met Professor Guust Nolet zeer gewaardeerd en met zijn vertrek naar Princeton verloor ik een inspirerend promotor. Ik ben Professor N.J. Vlaar erkentelijk voor het overnemen van deze taak.

Verder mijn dank aan Suzan van der Lee, Dolores Alsina, Henk Marquering en Harm Dorren, voor het kritisch lezen van verschillende delen van dit proefschrift.

Hanneke Paulssen, Bernard Dost, Michael Bostock and Nathalie Wajeman were always ready to lend an ear and a helping hand.

Ik ben Dee Pattynama veel dank verschuldigd voor haar hulp bij de laatste voorbereidingen voor dit proefschrift.

Verder wil ik alle collega's en studenten bedanken voor de prettige sfeer op het instituut, waardoor ik een zeer aangename tijd op de universiteit heb doorgebracht.

Jos wees mij de weg naar de geofysica, waarvoor ik hem nog steeds dankbaar ben.

Ine, je hebt me altijd gesteund en ervoor gezorgd dat ik de rest van de wereld niet vergat.

## Curriculum Vitae

- February 24 1965                      Born in Oldenzaal, The Netherlands
- 1977-1983                              Gymnasium- $\beta$ , Katholiek Veluws College, Apeldoorn, The Netherlands
- May 27 1988                            Doctoraal examen (drs. degree) geophysics, Utrecht University, The Netherlands
- 1988-1992                              Research assistant at the department of Theoretical Geophysics, Utrecht University, The Netherlands
- September-October 1992            Teaching assistant at the department of Theoretical Geophysics, Utrecht University, The Netherlands
- November 1992 - January 1994    Fulfillment of military service at the Physics Electronics Laboratory (FEL) of TNO, Den Haag.