Geophys. J. Int. (2003) **153,** 735–752

Inversion of shallow-seismic wavefields: II. Inferring subsurface properties from wavefield transforms

Thomas Forbriger*

Institut für Meteorologie und Geophysik, J. W. Goethe-Universität Frankfurt, Feldbergstraße 47, D-60323, *Frankfurt am Main, Germany.*
E-mail: Thomas.Forbriger@gpi.uni-karlsruhe.de

Accepted 2003 January 28. Received 2003 January 27; in original form 2002 September 10

SUMMARY

The problem of inferring subsurface properties from shallow-seismic data is solved by a two-stage scheme that fits the full wavefield by its synthetic counterpart. In a first stage (described in a companion paper) I derive Fourier–Bessel expansion coefficients for the recorded data through a wavefield transformation. The present paper describes the joint inversion of these coefficients together with *P*-wave arrival times to infer subsurface properties. In this way we exploit the full signal-content including the dispersion of higher-modes, leaky-modes, their true amplitudes, and, at least partly, body waves.

Owing to the multi-mode character of shallow-seismic field-data conventional techniques of dispersion analysis are not applicable. Since an initial model appropriate for inversion of full seismograms is rarely available in shallow seismics, the direct inversion of waveforms is not feasible. The wavefield transformation removes a remarkable amount of non-linearity from the data. In consequence the proposed method is robust even in the absence of *a priori* information. In distinction to the inversion of dispersion curves, it does not require the identification of normal-modes prior to inversion. The method performs well when applied to the multi-mode wavefields present in most shallow-seismic data sets. Compared to waveform fitting, it can be more efficient by about a factor of ten, because we need not evaluate the Bessel-expansion and therefore need less calculations of the forward problem.

Subsurface properties are derived for the two sets of field-data that were already presented in the first paper. One of them includes a pronounced low-velocity channel. For both we observe a remarkably good resolution of *S*-velocity down to the bedrock, which is found in 6 and 16 m depth, respectively. In both cases it would not be possible to infer the depth of bedrock from *P*-wave data alone. Synthetic seismograms calculated from the final model match the recorded waveforms surprisingly well, although no waveform fitting was applied. A subsequent waveform inversion becomes feasible with initial models taken from the results of this method.

Finally it is shown by example, that conventional techniques of dispersion-curve fitting are likely to give misleading results when applied to our field-data.

Key words: full wavefield inversion, joint inversion, low-velocity channel, near-field, Rayleigh waves, shallow seismics.

1 INTRODUCTION

1.1 Application of shallow surface-wave inversion

The shear wave velocity in the subsurface is a material property of interest for many applications like building site characterization, ground-water studies, shear strength estimates, or for providing static corrections in shear-wave surveys. Rayleigh waves provide direct access to shear wave velocity, are easy to excite and record and offer an excellent signal-to-noise ratio. In contrast to refraction seismics, surface wave studies are not limited to media with seismic velocities increasing with depth. Therefore they are well applicable in geotechnical tests of the subsurface of pavements.

Surface-wave studies are rather frequent in civil-engineering literature (e.g. Jones 1958; Nazarian 1984; Gucunski & Woods 1991; Tokimatsu *et al.* 1992; Matthews *et al.* 1996). Many of these studies apply a method called SASW (spectral analysis of surface waves) by

*Now at: Black Forest Observatory (BFO), Heubach 206, D-77709 Wolfach, Germany.

Stokoe & Nazarian (1983). An 'Annotated Bibliography on SASW' compiled by Hiltunen and Gucunski lists 41 publications in the years from 1982 to 1993 (Oelsner, private communication, 1999). 27 of them were published in the period from 1988 to 1992. And 20 studies deal with the subsurface of pavements. These studies adopt techniques originally developed for teleseismic investigations. Most of them neglect typical properties of shallow-seismic wavefields like mode interference, dominance of higher modes, osculation points, and strong leaky-modes. Complications arising from this are discussed in Paper I (Forbriger 2003, this issue).

1.2 The concept of the proposed inversion scheme

The proposed two-stage inversion infers subsurface properties from a wavefield transform instead of dispersion curves or waveforms. I define the $(\omega,\ p)$ transform $G_\eta(\omega,\ p)$ of the recorded wavefield by the expansion

$$u_\eta(t,r) = \int_{-\infty}^{+\infty} \int_0^{+\infty} G_\eta(\omega,\ p) J_\eta(\omega p r) p\, dp\, e^{-i\omega t} \frac{d\omega}{2\pi}, \qquad (1)$$

where $\eta = 0$ for the vertical component and $\eta = 1$ for the radial component, respectively (Paper I). $u_\eta(t,r)$ is the corresponding seismogram at offset r and J_η is the Bessel function of the first kind. The $(\omega,\ p)$ transform is a complete representation of the recorded wavefield. It allows one to exploit the full signal-content including the dispersion of higher-modes, leaky-modes, and their true amplitudes. It avoids to deal with dispersion in terms of normal modes and thus any complications arising from mode identification. The relation between the data and subsurface parameters is significantly less non-linear than it is for waveform data. Therefore the inversion is robust even in the absence of *a priori* information and it is less likely to be trapped in local minima of the objective function. Further, this approach is potentially more efficient than waveform fitting, as is substantiated in Paper I.

In addition to the $(\omega,\ p)$ spectrum we use onset-times of refracted *P*-waves in a joint inversion. Basically the body waves are also present in the wavefield transform. However, since refracted waves are typically of small amplitude and large apparent velocity, they can be severely disturbed in the spectrum by surface waves that are either excited by noise sources or scattered by heterogeneity and appear later in the waveforms but with large apparent phase-velocities in the geophone line. The onset time of refracted *P*-waves is a unique information and is easy to extract. It gives additional constraints on the *P*-velocity structure that may be only weakly constrained by the dispersion of surface waves.

We fit both the $(\omega,\ p)$ transform and the *P*-wave arrivals by synthetic predictions. The calculation of refracted wave traveltimes is straightforward for horizontally layered media (Telford *et al.* 1990, section 4.3.3). Fourier–Bessel expansion coefficients for the impulse response of the medium as defined by eq. (1) can be taken from standard algorithms for the calculation of synthetic seismograms in plane, laterally homogeneous media. These algorithms typically expand the wavefield into cylindrical harmonics. A reflectivity code written by Ungerer (1990) was used for the examples below.

1.3 The structure of this paper

In Section 2 I describe particularities of the iterative least squares problem where they go beyond textbook subjects. This concerns control parameters like data weights, model damping, model parametrization, and reasonable parameter ranges. I also have to discuss the treatment of the source wavelet. It is an essential part of the data that cannot be described by a simple δ-function. After describing the inversion itself, I add a resolution analysis that tests how well constrained the inferred parameters are.

Two quite different field data examples are shown in Section 3. The first reveals a smooth but strong velocity variation in the subsurface. The wavefield is dominated by a higher mode. The second data set was recorded on a hard pitch. The site has a strong low-velocity channel due to the hard top asphalt layer. This results in very specific dispersion characteristics (Paper I). The inversion is able to resolve v_s in the channel. In both cases synthetic waveforms are calculated from the final model. They match the recorded data already quite well. Remarkably, this is achieved without waveform fitting.

Finally, in Section 4, I discuss systematic problems that arise when conventional techniques are applied to the example data sets. They are likely to result in misinterpretations.

2 INVERSION SCHEME

2.1 Setting up the inverse problem

To outline the elements of the inverse problem I define the penalty function

$$E_n^2 = |\mathbf{W}(\mathbf{d} - \mathbf{s}(\mathbf{m}_n) - \mathbf{D}\delta\mathbf{m}_n)|^2 + \lambda|\mathbf{S}\delta\mathbf{m}_n|^2 \qquad (2)$$

of the *n*th iteration. The vector $\mathbf{m}_n$ contains the initial values of the inversion parameters for the *n*th iteration. This may include all parameters describing the subsurface properties (seismic velocities, depth of first order discontinuities, *Q*-values and density) or just a subset of them. The elements of the vector $\mathbf{d}$ are the data values, which here are both the real *P*-wave arrival times and the complex Fourier–Bessel coefficients of the recorded wavefield. The vector $\mathbf{s}(\mathbf{m}_n)$ denotes the corresponding synthetic predictions, based on model $\mathbf{m}_n$. The elements

$$D_{kl} = \left.\frac{\partial s_k}{\partial m_l}\right|_{\mathbf{m}_n} \qquad (3)$$

of the matrix of partial derivatives $\mathbf{D}$ are complex valued like the wavefield expansion coefficients, and have to be recalculated after each model-update. The matrix $\mathbf{W}$ contains real weight factors for the data and the matrix $\mathbf{S}$, in the damping term, applies weights to the parameter variations. The factor λ allows to scale the damping, but usually $\lambda = 1$.

The least squares condition $E_n^2 \stackrel{!}{=} \min$ leads to the system of linear equations

$$\Re\{(\mathbf{D}^\dagger\mathbf{W}^{\mathrm{T}}\mathbf{W}\mathbf{D}) + \lambda\mathbf{S}^{\mathrm{T}}\mathbf{S}\}\delta\mathbf{m}_n = \Re\left\{\mathbf{D}^\dagger\mathbf{W}^{\mathrm{T}}\mathbf{W}(\mathbf{d} - \mathbf{s}(\mathbf{m}_n))\right\}, \qquad (4)$$

where $\Re\{z\}$ means the real part of the complex value z, $\mathbf{D}^\dagger$ is the hermitian conjugate of $\mathbf{D}$, and $\mathbf{W}^{\mathrm{T}}$ is the transpose of $\mathbf{W}$. The next improved estimate of model parameters follows from the solution of eq. (4) by

$$\mathbf{m}_{n+1} = \mathbf{m}_n + \delta\mathbf{m}_n. \qquad (5)$$

The first initial model $\mathbf{m}_1$ still must be found by trial and error.

By solving eqs (4) and (5) iteratively we systematically improve the model parameters. The iteration is stopped when the relative change of E^2 or $\mathbf{m}$ falls below a case-specific limit (i.e. when the operating seismologist is satisfied) or when $\mathbf{m}$ leaves the range of reasonable values.

2.2 Data weights and misfit

In its simplest form $\mathbf{W} = \mathrm{diag}(W_k)$ is a diagonal matrix, where $W_k = 1/\sigma_k$ accounts for the tolerance σ_k of the data d_k. By this all residuals in eq. (2) become dimensionless relative values, regardless of their physical nature. We then may easily combine traveltime data and expansion coefficients. Aliasing should be considered when defining the W_k (Paper I). Practically they allow to suppress the misleading influence of aliasing in the (ω, p) coefficients.

Since the number of (ω, p) coefficients is typically much larger than that of first breaks, the former will dominate in the inversion. If this is not desired, extra factors have to be applied to the W_k to balance between both data sets.

It is advisable to normalize the W_k according to the total number of the data values so that a misfit

$$\chi^2(\mathbf{m}) = |\mathbf{W}(\mathbf{d} - \mathbf{s}(\mathbf{m}))|^2 \tag{6}$$

equal to one means that the data are in the average explained by the subsurface model $\mathbf{m}$ within their tolerance. In cases where the physical meaning of the W_k in the sense of data tolerance is questionable, we simply normalize

$$\chi^2(\mathbf{m}) = \frac{|\mathbf{W}(\mathbf{d} - \mathbf{s}(\mathbf{m}))|^2}{|\mathbf{W}\mathbf{d}^2|} \tag{7}$$

by the weighted mean square of the data. I prefer to achieve this by choosing the W_k appropriately for $|\mathbf{W}\,\mathbf{d}|^2 = 1$, thus making the denominator in eq. (7) obsolete. For $\mathbf{s} = 0$ we then obtain $\chi^2 = 1$. However, $\chi^2 = 1$ practically means some success in fitting the data, because we will ensure $|\mathbf{s}| \approx |\mathbf{d}|$ in the inversion process.

χ^2 measures the portion of the signal energy that is not fitted. Although this measure is essential to least-squares fitting, its exact value may be of little objective significance. It depends on highly subjective decisions like the relative weight given to the *P*-wave data relative to the (ω, p) spectrum, the section of the spectrum used for the inversion, and the distribution of the W_k. Besides, the reduction in misfit that actually can be achieved depends on the model parametrization. For this reason, I refer to the graphical comparison of data with synthetics and especially to the waveforms when assessing the finally achieved fit, rather than reducing this information to the scalar χ^2.

2.3 Damping

Apart from the misfit, the penalty function (2) contains an extra term $|\mathbf{S}\,\delta\mathbf{m}|^2$. This is used to damp the iterative inversion. It ensures the regularity of the system (4). When the elements of $\mathbf{S}$ are well chosen, it keeps the model modifications $\delta\mathbf{m}$ in a range, where the linearization

$$\mathbf{s}(\mathbf{m} + \delta\mathbf{m}) \approx \mathbf{s}(\mathbf{m}) + \mathbf{D}\delta\mathbf{m}, \tag{8}$$

inherent in eq. (4), is a useful approximation.

The matrix $\mathbf{S}$ applies weights to the model changes $\delta\mathbf{m}$. In its simplest form it is a diagonal matrix

$$\mathbf{S} = \mathrm{diag}\left(\frac{1}{r_k}\right), \tag{9}$$

where r_k may be regarded as a kind of search range for the parameter m_k. That is the range $\delta m_k = \pm r_k$, where reasonable improvements are expected from the linear approximation (i.e. where eq. 8 holds). Since the damping term is added to the misfit in eq. (2), the r_k should contain appropriate normalization to balance between both terms. Scaling all r_k with $\sqrt{N}$, where N is the number of inversion parameters m_k, results in a damping term equal to one if the full search ranges are exploited in the average.

Appropriate search ranges are found by trial and error. For suitable r_k the misfit $\chi^2(\mathbf{m} + \delta\mathbf{m})$ should decrease with increasing λ, where $\delta\mathbf{m}$ is the solution of eq. (4) and thus a function of λ. If not, the model variation $\delta\mathbf{m}$ exceeds the range of the linear approximation (8) and the search ranges must be chosen smaller. For the examples that are discussed below suitable ad hoc values for r_k have been found to be 100 m s^{-1} for the seismic velocities and 0.1 m for the depth of discontinuities.

The damping term introduced here limits the model variation within one iteration step. The total model variation relative to the first initial model $\mathbf{m}_1$ still may be arbitrary large. While the terms present in eq. (2) are a must, it might be helpful to add additional damping terms relating the model parameter to a reference or to introduce empirical v_p–v_s-correlations or a smoothing constraint (Forbriger 2001). In many cases this can be problematic due to the notorious lack of *a priori* data in shallow-seismic studies. Nevertheless, it stabilizes the inversion, and can replace manual interventions in the iterative process by weak constraints.

2.4 Partial derivatives and model parametrization

If partial derivatives (3) are not available in analytical form, we have to approximate them through finite differences. We then need one extra calculation of synthetics $\mathbf{s}(\mathbf{m})$ per inversion parameter m_k. This is the numerically most expensive part of the whole process. We therefore will try to keep the number of model parameters small. Notice that we will need much more iterations to reach the final model than is usual in teleseismic studies, where initial models typically differ only by a few percent from the final result.

Experience shows that in some cases a stack of only a few homogeneous layers is inappropriate to describe the observed wave propagation. In the case of Bietigheim (discussed below) we need a non-linear continuous variation of the seismic velocities with depth. Bachrach *et al.* (2000) theoretically predict an increase of v_p and v_s with depth to the power of 1/6 in sand and confirm this by experimental observations.

In the examples below I used an expansion of the model functions into second order polynomials. This is a good compromise between the demands to use only few inversion parameters and to describe smooth variations with depth. However, further experiments with waveform-fitting for the example Bietigheim, starting from the final model presented below, resulted in a significantly improved fit, when allowing parameter variations that depart from second order polynomials in the top few metres. The associated small changes of the velocity model might be regarded as insignificant. But, since the partial derivatives are model-dependent, this might give hope to exploit some of the weak constraints on density present in the data, when allowing more general model variations for the seismic velocities.

In the examples below the inversion parameters were polynomial coefficients, while the applied reflectivity algorithm uses a stack of homogeneous layers. Therefore the forward calculation of synthetic data includes the calculation of a layer-approximation to the continuous model-functions.

2.5 Ranges of reasonable parameter values

The materials met in the shallow depth range include unconsolidated sediments and solid bedrock. At interfaces between these two, the

seismic velocities may change by a factor of five or more. For comparison, prominent global discontinuities are much smaller (e.g. 20 per cent at the Moho). Within the unconsolidated sediments we find material ranging from muddy loam to dry sand. Owing to the lack of *a priori* information, we have to expect the inversion results within a wide parameter range. Since the least-squares criterion itself does not ensure the inversion results to be physically feasible, we have to check them to lie within reasonable ranges.

2.5.1 Velocities and Poisson's ratio

Observed *P*-velocities range from 40 m s^{-1} in dry sand (Bachrach *et al.* 1998) to a few km s^{-1} in solid bedrock. The *S*-velocities vary over a similar range but are linked to v_p. The v_p/v_s ratio is reasonably expected between 1.56 (i.e. a Poisson number of $\nu = 0.15$) and values larger than 10 (i.e. $\nu > 0.495$). $\nu \approx 0.15$ was observed by Bachrach *et al.* (2000) in dry sand and in our example Bietigheim. $\nu > 0.495$ approaches the case of a liquid ($v_s = 0$ m s^{-1}, $\nu = 0.5$). Roth & Holliger (2000) report values of $v_p/v_s = 16$ ($\nu = 0.498$). Seismic material properties of shallow sediments were also studied intensively by Stümpel *et al.* (1984). They report v_p/v_s ratios up to 9 ($\nu = 0.494$) in boulder clay and variations from 1.41 ($\nu = 0$) to 4 ($\nu = 0.47$) in dry and partially saturated sand.

2.5.2 Dissipation

Attenuation generally may be strong ($Q < 10$) in unconsolidated sediments. Positive seismic velocities and *Q*-values do not ensure positive elastic moduli, which are essential to thermodynamic stability of the material. Therefore it is necessary to calculate the complex bulk and shear moduli at each iteration and to check them for positive real parts and negative imaginary parts (due to the sign convention of the Fourier transformation used in eq. 1). Then all other elastic moduli satisfy this condition too.

The properties of dissipative material must be frequency dependent to be causal. In fact sediments show not only frequency-dependent moduli but also frequency-dependent *Q* in the seismic frequency-range (e.g. Spencer 1981; Burkhardt *et al.* 1992). It is, however, almost hopeless to infer additional information on the intrinsic frequency-dependency. We use velocities and *Q*-values constant with frequency in the inversion, because the depth variation of seismic velocities is much larger than the intrinsic frequency-variation (even for small *Q*).

Stronger attenuation leads to broader surface-wave peaks in the (ω, p) domain. However, also spatial tapering influences the form of the resonance peaks. For this reason I regard the resulting *Q*-models as rather qualitative estimates. Comparison of recorded waveforms with synthetic seismograms calculated from the subsurface models shows that attenuation is typically underestimated. This is due to initial *Q*-models that are still chosen too conservative.

Global reference models like PREM (Dziewonski & Anderson 1981) indicate that dissipation may be mostly due to shear deformation. If this applies to shallow media too, it would be more appropriate to search for Q_μ rather than Q_p and Q_s, and to keep Q_κ large and fixed.

2.5.3 Density

In general surface waves are sensitive to variations in density too. However, the sensitivity to density is significantly smaller than to seismic velocity. Therefore we do not vary the values taken for the initial model except to test for vanishing constraints on density. Ad hoc values for the initial model are available from standard references (e.g. Schön 1998).

2.6 The source time-function

The (ω, p) coefficients

$$d_K = G(\omega_k, p_l) \tag{10a}$$

of the wavefield are calculated from the recorded waveforms with respect to eq. (1). Here $K = \{k, l\}$ denotes unique combinations of k and l. The corresponding synthetic predictions, which shall fit the data, are

$$s_K = S(\omega_k) G_\delta(\omega_k, p_l), \tag{10b}$$

where $G_\delta(\omega_k, p_l)$ is the impulse response of the subsurface model and $S(\omega_k)$ are Fourier coefficients of the impulsive response of the experimental setup (further called 'field-device response'). The latter contains both the known contributions like the geophone response or the anti-aliasing filter response and the unknown waveform of the source (e.g. hammer impact). Since the time constant of the source process lies within the recorded period band, we cannot simply approximate it by a delta-impulse function. In order to fit the data through synthetics we have either to remove the influence of $S(\omega_k)$ or to determine it from the data by inversion.

The field-device response coefficients $S(\omega_k)$ could be regarded as an extra set of unknowns or model parameters in the inversion. Then they would be determined in the inversion process along with the subsurface parameters. I prefer to treat them separately from the inversion parameters. Two different approaches are discussed below. The first removes the $S(\omega_k)$ by normalization of data and synthetics. The second determines appropriate $S(\omega_k)$ within each forward calculation of synthetic predictions s_K. Thus as much as possible is removed from the residuals between data and synthetics before they enter eq. (2). We use only the unexplained rest to infer subsurface structure. I regard this as the most conservative approach.

In a similar way the traveltime may contain a systematic delay due to the trigger mechanism and filter delays. This can be removed by normalization or can be determined from the data as well, similar to the handling of (ω, p) coefficients explained below.

2.6.1 Normalizing the (ω, p) coefficients

Alternatively to eqs (10) we may define the data

$$d_K = \frac{G(\omega_k, p_l)}{\sum_j W_{\{k,j\}} G(\omega_k, p_j)} \tag{11a}$$

and the synthetic predictions

$$s_K = \frac{G_\delta(\omega_k, p_l)}{\sum_j W_{\{k,j\}} G_\delta(\omega_k, p_j)}, \tag{11b}$$

with $W_{\{k,j\}} = W_K$ being the appropriate diagonal elements of the weight matrix $\mathbf{W}$ in eq. (2).

This way all (ω, p) coefficients are normalized to a weighted mean of 1 in a sum over all slowness samples at each frequency. The influence of the field-device response is thus removed from both, data and synthetics. The latter will fit the former if appropriate subsurface parameters are used to calculate $G_\delta(\omega_k, p_l)$. The (ω, p) coefficients may be small due to destructive interference of waves or may compensate in the sum. Hence a lower bound should be applied to the denominator of eqs (11) to avoid numerical overflow.

2.6.2 Determining the field-device response

The subsurface parameters and the source time-function have rather independent influence to the predictions (10b). While the subsurface parameters influence the location, width, and amplitude of individual maxima of $G_\delta(\omega_k, p_l)$ in the (ω, p) plane, the coefficients $S(\omega_k)$ of the field-device response will affect the amplitude and phase of all coefficients at a given frequency in the same way. The (ω, p) coefficients at a single phase-slowness can be explained by an appropriately chosen $S(\omega_k)$ alone. If we, however, demand a good fit for all p_l at once we can determine subsurface parameters and field-device response independently.

Field-device response coefficients

$$S(\omega_k) = \frac{\sum_l W^2_{\{k,l\}} G^*_\delta(\omega_k, p_l) G(\omega_k, p_l)}{\sum_l W^2_{\{k,l\}} G^*_\delta(\omega_k, p_l) G_\delta(\omega_k, p_l)} \tag{12}$$

fit the data in a least-squares sense. Here G^*_δ is the complex conjugate of G_δ. Again it is advisable to apply a lower bound to the denominator.

The true force time-function of the hammer can be estimated by fitting a filtered trial function to $S(\omega_k)$. Examples are discussed below.

2.6.3 When to use which of both alternatives

At an early stage of the iterative inversion the $S(\omega_k)$ determined from eq. (12) are completely meaningless since $G_\delta(\omega_k, p_l)$ is still far from describing the observed wave propagation. I therefore propose to start the inversion with the normalization of (ω, p) coefficients. This approach also removes the field-device amplitude spectrum and thus allows an equal contribution of all frequencies to the inversion result.

At that early stage it also might be advisable to use the real absolute value of $G_\delta(\omega_k, p_l)$ and $G(\omega_k, p_l)$ rather than the complex values. The varying phase may destabilize the inversion as long as the dispersion characteristics of the data and synthetics are in bad misfit. As well it can be helpful to use the logarithm of the amplitude. This allows modes with smaller amplitudes in the wavefield-transform to contribute stronger to the resulting subsurface model.

It is not necessary to actually determine the $S(\omega_k)$ until we really want to infer the time-function or calculate synthetic seismograms for comparison with recorded waveforms.

When inverting only the modulus of $G(\omega_k, p_l)$, it can be helpful to calculate a real amplitude factor by eq. (12) rather than using normalization. This will lead to vanishing amplitudes of synthetic predictions and partial derivatives at frequencies where they do not match the observed data at all. Thus implicitly only those spectral intervals influence the inversion where the predictions are already appropriate. This stabilizes the whole procedure. With ongoing iteration larger (ω, p) intervals will contribute.

2.7 The inversion process

2.7.1 The initial model

An initial model for v_p is most easily found by standard refraction inversion. As a first guess the v_s model may be taken from $v_s = v_p/\sqrt{3}$. Density is taken from standard references and Q-values are simply guessed. The latter should be taken rather too small than too large, since large Q results in narrow maxima of large amplitude in the (ω, p) spectrum. For broad maxima the iterative inversion becomes more robust at an early stage where data and synthetics hardly match.

Usually the initial v_s-model is far from explaining the (ω, p) spectrum of the observed wavefield. A suitable initial model for the least-squares inversion must be found by trial and error. If the surface wave maxima in the predicted (ω, p) spectrum are at phase slowness values smaller than in the recorded wavefield, v_s-values in the model have to be decreased and vice versa. If the transition to larger phase slowness values of a dispersion curve appears at frequencies larger than in the observed data, the depth of model discontinuities has to be increased and vice versa. This manual procedure must be continued until the predicted and the observed maxima touch at least at one frequency.

2.7.2 Twofold iteration

The actual least-squares inversion involves a twofold iteration. The automatic least squares inversion must be iterated due to the non-linearity of the problem. This process of improving inversion parameters cannot be done manually due to the multidimensionality of the parameter space and trade-off between parameters. However, the automatic inversion may produce unreasonable results in some regions of the model due to noise in the data, lack of resolution, inappropriate parametrization, and lack of a reliable reference model to which we could constrain the inversion. For this reason we have to intervene manually by explicitly changing model values, removing or adding model parameters from/to the set of inversion parameters, changing additional constraints or weighting parameters, etc. Then the automatic inversion is started again, now with another, already improved initial model.

This scheme is iterated until the automatic inversion no longer produces a significant misfit reduction. Apart from being unavoidable, the frequent manual interventions help the operator to gain an intuitive knowledge about the resolution capabilities of the particular data set. Besides, it is useful to check for extremes of possible parameter values like the maximum depth of the halfspace.

Each sequence of automatic iterations may include 10 to 50 steps. And the whole process may include 10 to 20 manual interventions depending on how many different model alternatives should be tested. On a modern Personal-Computer this can conveniently be done in interactive mode.

2.8 Resolution analysis ('Rubberband test')

Resolution analysis for non-linear inversion problems can hardly be accomplished by a single test. At many places in the parameter space we might find alternative models that resemble the inferred model in the quality of the data-fit obtained from them. To test for this, the manual interventions in the inversion procedure are essential. Local, quantitative resolution analysis is based on the matrix of partial derivatives. I supplement the inversion results, given in the examples below, with a simple (but instructive) linear test. It gives a local estimate of constrainedness by linear extrapolation of the predicted data. The concept of this test is due to Wielandt (private communication, 1998). A similar, more elaborate discussion is given by Jackson (1976).

We start from an optimized model $\mathbf{m}_{\rm opt}$, i.e. a minimum of E^2. Varying the Lth model parameter m_L by Δm_L will increase the misfit between predicted and recorded data. At the same time we may add changes δm_l to all other model parameters to keep the increase in the misfit as small as possible. This accounts for the trade-off between model parameters. How large is the allowed model variation in this sense, when allowing a maximum increase of $\Delta\chi^2_{\rm max}$ in the misfit? I

will ask this question in turn for every model parameter. This test is easily implemented in the standard least-squares algorithm as will be explained.

The test can be visualized as follows: I seize one model-parameter curve at any depth. Then I pull it to one direction (e.g. increasing velocity or decreasing the depth of a discontinuity). The curve will bend, like a rubber band, but with different shape. The rubber band bends in a way that keeps the potential energy small. The model curve bends in a way that keeps $\Delta\chi^2$ small. The weaker the constraints, the more bending is possible until a threshold $\Delta\chi^2_{\max}$ is reached.

Rather than discussing only the misfit I use the penalty function

$$E^2(\delta\mathbf{m}+\mathbf{m}_{\rm opt}) = |\mathbf{W}(\mathbf{d}-\mathbf{s}(\mathbf{m}_{\rm opt})-\mathbf{D}\delta\mathbf{m})|^2 + \lambda|\mathbf{S}\delta\mathbf{m}|^2. \quad (13)$$

Here $\mathbf{m}_{\rm opt}$ is the inversion result, the 'optimum' model, a minimum of E^2. The partial derivatives $\mathbf{D}$ are calculated for this model. The factor λ allows to study the influence of the damping conditions on the inversion result. A test for $\Delta\chi^2$ is accomplished by setting $\lambda = 0$.

We now study the increase

$$\Delta E^2(\Delta m_L) = E^2(\delta\mathbf{m}+\mathbf{m}_{\rm opt}) - E^2(\mathbf{m}_{\rm opt}) \quad (14)$$

of E^2, where we set the Lth element of $\delta\mathbf{m}$

$$(\delta\mathbf{m})_L \stackrel{!}{=} \Delta m_L \quad (15)$$

and optimize all other elements in $\delta\mathbf{m}$ to keep ΔE^2 small. Abbreviating $\mathbf{M} = \mathbf{D}^\dagger\mathbf{W}^{\rm T}\mathbf{W}\,\mathbf{D} + \lambda\mathbf{S}^{\rm T}\mathbf{S}$ we obtain

$$\begin{aligned}\Delta E^2(\Delta m_L) &= \delta\mathbf{m}\mathbf{M}\delta\mathbf{m}\\ &= \sum_{l\neq L, n\neq L}\delta m_l M_{ln}\delta m_n + \sum_{l\neq L}\delta m_l \Delta m_L (M_{lL}+M_{Ll})\\ &\quad + \Delta m_L M_{LL}\Delta m_L. \end{aligned} \quad (16)$$

Since $\mathbf{m}_{\rm opt}$ itself is an extremum of E^2, $\Delta E^2(\Delta m_L)$ has only terms quadratic in $\delta\mathbf{m}$. The sums are formed for all values of l and n, except for $l = L$ and $n = L$. Keeping ΔE^2 small by optimizing all δm_l with $l \neq L$ is a least squares criterion equivalent to

$$\frac{\partial\Delta E^2}{\partial\delta m_l} = 0 \quad \text{for all } l \neq L \quad (17)$$

which is

$$\sum_{n\neq L} R_{ln}\delta m_n + R_{lL}\Delta m_L = 0 \quad (18)$$

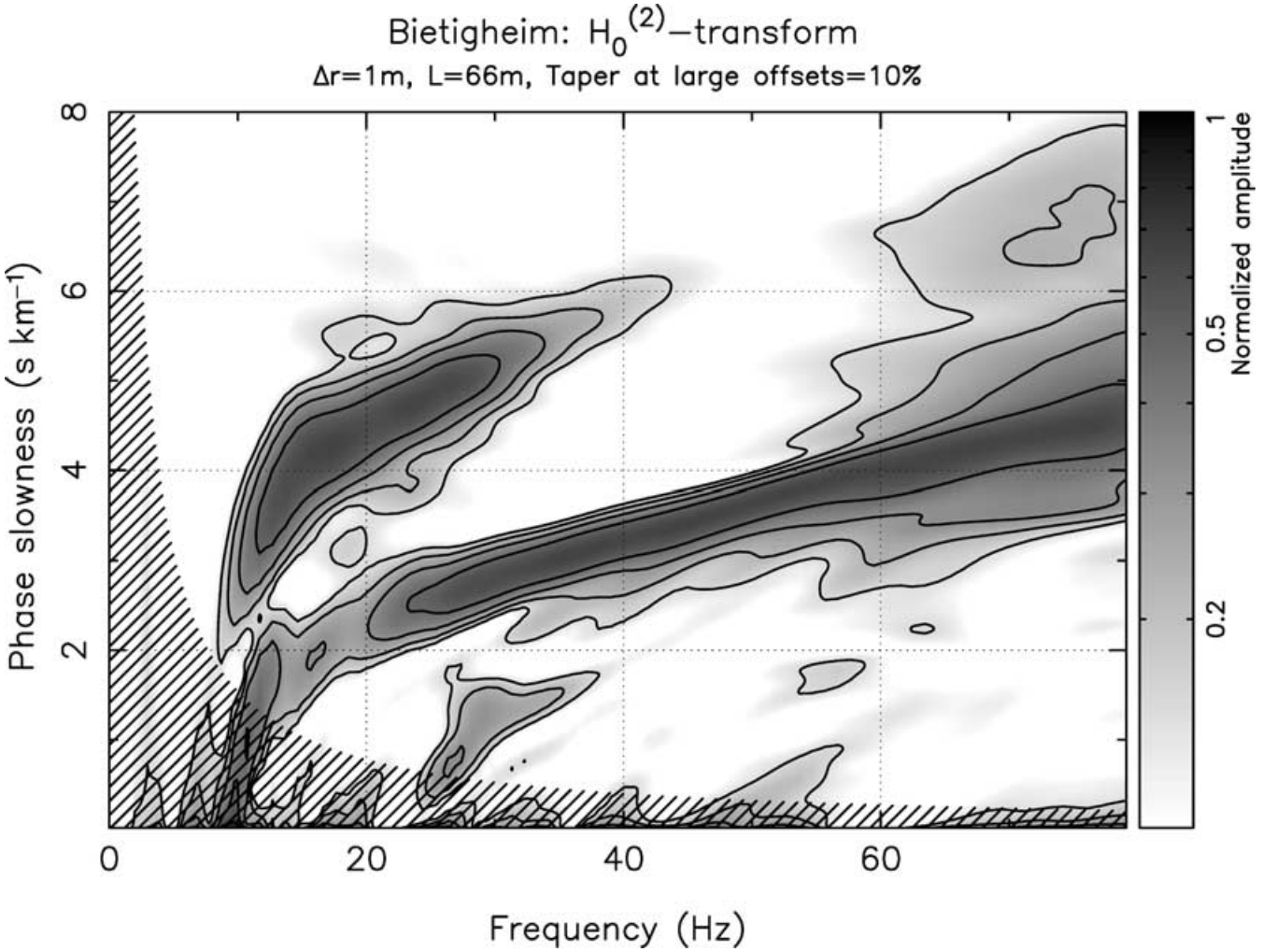

Figure 1. Data set Bietigheim: (ω, p) spectrum of the recorded data. The grey-scale gives the modulus of the complex wavefield coefficients that are calculated with the modified Hankel transformation (Paper I). The most prominent signals are the fundamental Rayleigh mode between 10 and 30 Hz and the first higher mode that dominates at frequencies larger than 30 Hz. The hatched area gives a measure of the maximum resolution possible for a spread length of 66 m. Aliasing is avoided by a dense effective geophone interval of 1 m in the combined data set.

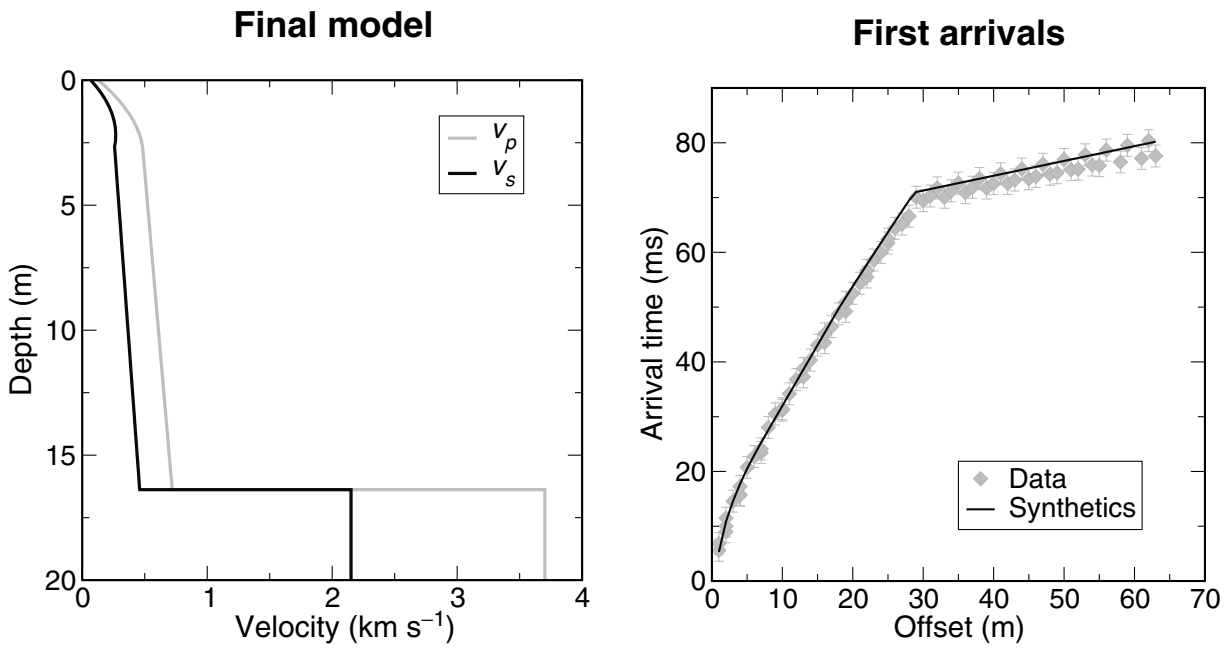

Figure 2. Data set Bietigheim: Final model and traveltime fit. The left graph gives the variation of seismic velocities with depth as inferred by a joint inversion of the (ω, p) spectrum (Fig. 1) and the *P*-wave first arrivals. The right graph gives the picked *P*-wave arrival times with error bars and the traveltimes predicted by the final model.

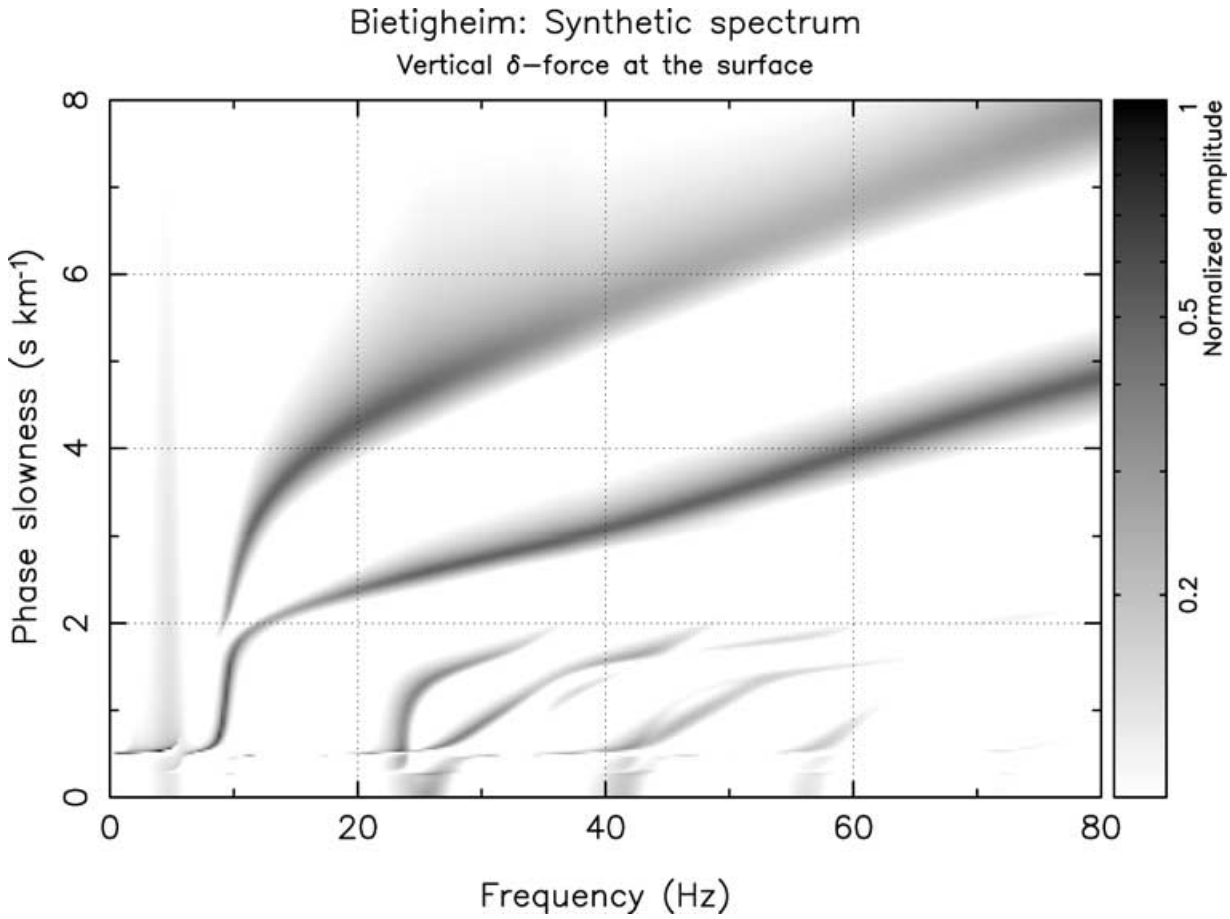

Figure 3. Data set Bietigheim: Spectrum of the synthetic Green's function $G_\delta(\omega, p)$ for the final model (Fig. 2). It fits the (ω, p) spectrum of the data (Fig. 1) in a least-squares sense. The fundamental mode and the first higher mode are predicted correctly. Like in the data, the first higher mode dominates at frequencies larger than 30 Hz. Whether the feature in the data at 25 Hz and 1 s/km can be explained by the second and third higher mode remains questionable due to the limited data quality.

for all $l \neq L$ with $\mathbf{R} = \Re(\mathbf{M})$. We obtain a numerical solution for the δm_n by setting

$$\Delta m_L = 1. \tag{19}$$

Eqs (18) and (19) can be combined to the system of linear equations

$$\mathbf{R}'\delta\mathbf{m}_L = \mathbf{e}_L, \tag{20}$$

where $\mathbf{R}'$ is derived from $\mathbf{R}$ by setting the Lth row to zero except the Lth element which is set to one. The vector $\mathbf{e}_L$ has only elements of zero, except the Lth element, which is one. The solution $\delta\mathbf{m}_L$ of eq. (20) is the model variation with least ΔE^2 when setting the model parameter $\Delta m_L = 1$. From this we calculate

$$\Delta\mathbf{m}_{L\max} = \sqrt{\frac{\Delta E^2_{\max}}{\delta\mathbf{m}_L\mathbf{M}\delta\mathbf{m}_L}}\,\delta\mathbf{m}_L \tag{21}$$

which is the total model variation for a maximum variation of model parameter L while allowing for an increase of E^2 by $\Delta E^2_{\max}$. By comparison of eqs (4) and (20) we find that this resolution analysis can be carried out with the inversion algorithm by simply modifying the Lth row of the system matrix and the inhomgeneous term in eq. (4).

In the examples below, I will display $\mathbf{m}_{\rm opt} \pm \Delta\mathbf{m}_{L\max}$ rather than $\pm\Delta\mathbf{m}_{L\max}$ alone, because this graphically links the test results to the final model. It must be emphasized however, that $\mathbf{m}_{\rm opt} \pm \Delta\mathbf{m}_{L\max}$, although discussed in model space, are not valid alternative models. The test gives a *local estimate of constrainedness by linear extrapolation*. In the case of model parameters being almost unconstrained by the data, $\mathbf{m}_{\rm opt} \pm \Delta\mathbf{m}_{L\max}$ may easily leave the range of reasonable parameter values and may even have negative values. The change ΔE^2 calculated by exact evaluation of $\mathbf{s}$ $(\mathbf{m}_{\rm opt} \pm \Delta\mathbf{m}_{L\max})$ may be larger or smaller than $\Delta E^2_{\max}$ in any case.

Since $\mathbf{m}_{\rm opt}$ is a minimum of E^2, all constant terms and terms linear in $\delta\mathbf{m}$ cancel out in eq. (14). The terms remaining in eq. (16) are quadratic in $\delta\mathbf{m}$ and independent of the data $\mathbf{d}$. ΔE^2 measures only the increase of E^2 relative to the best-fit synthetics, regardless of the total fit. Only ΔE^2 in relation to the constant terms in eq. (14) gives a measure relative to the data and I will specify $\Delta E^2_{\max}/E^2(\mathbf{m}_{\rm opt})$ in the examples below.

Notice that $\sqrt{\Delta E^2_{\max}}$ in eq. (21) serves as a scaling factor for the plot of $\mathbf{m}_{\rm opt} \pm \Delta\mathbf{m}_{L\max}$ in the first place. We will tend to choose small values to keep the less constrained parameters within the plot range. Furthermore, all remarks on χ^2 in Section 2.2 apply to ΔE^2, too. Eq. (21) implicitly depends on subjective decisions like the distribution of data weights or the range of (ω, p) coefficients used. Parts of the data that are badly fitted by $\mathbf{m}_{\rm opt}$ may contribute an apparent resolution just because $\Delta\mathbf{m}_{L\max}$ has an influence on the synthetics. Thus along with the formal resolution analysis the source of observed constraints must be discussed as is done in the examples below. In that context a look at the partial derivatives can be very instructive.

3 FIELD-DATA EXAMPLES

I already discussed two data sets in Paper I, illustrating the properties of the wavefield transform. They are examples for typical classes of subsurface structure. Here I continue with an inversion leading to a model for the subsurface of each of the discussed sites.

3.1 Data set Bietigheim

The (ω, p) spectrum of the data set, as discussed in Paper I, is shown in Fig. 1. By joint inversion of the (ω, p) coefficients and the first arrivals (Fig. 2, right) the subsurface model given in Fig. 2 (left) was obtained. The arrival times predicted by the model fit the observations. A spectrum of the Green's function $G_\delta(\omega, p)$ for this model is shown in Fig. 3. It reproduces the Rayleigh-wave dispersion of the observed wavefield as well as the amplitude excitation. The first higher mode dominates at frequencies greater than 30 Hz.

I distinguish three different regions in the model. The uppermost 2.5 m have small velocities with significant depth-dependent variation, here parametrized by second order polynomials. The material here is weathered loam. A linear gradient is not sufficient to explain the almost linear $p(\omega)$ relation of the dispersion above 20 Hz. At least a variation of second order is essential to produce this feature. Bachrach *et al.* (2000) show a similar strong variation of seismic velocities in the very shallow subsurface. They explain their observation by the pressure dependency of effective elastic moduli in sand. This may be a feature common to all unconsolidated sediments in the shallow region.

From 2 m to 16 m depth we observe intermediate velocities with a slight gradient. This region is still characterized by unconsolidated sediments (mostly loam and gravel). At a discontinuity in about 16 m depth the velocities jump to substantially higher values of a few km s^{-1}. This is the bedrock (dolomitic limestone). We have no resolution below that depth and thus define the bedrock as a homogeneous halfspace.

Table 1 gives a drill log that was taken at a distance of a few decametres from the seismic profile. We can identify the main features of the final model in the drilling result. A driving test at the seismic location could not be carried down to a depth larger than 9 m for technical reasons. It did not reach the bedrock, but confirmed the drill log in the upper part.

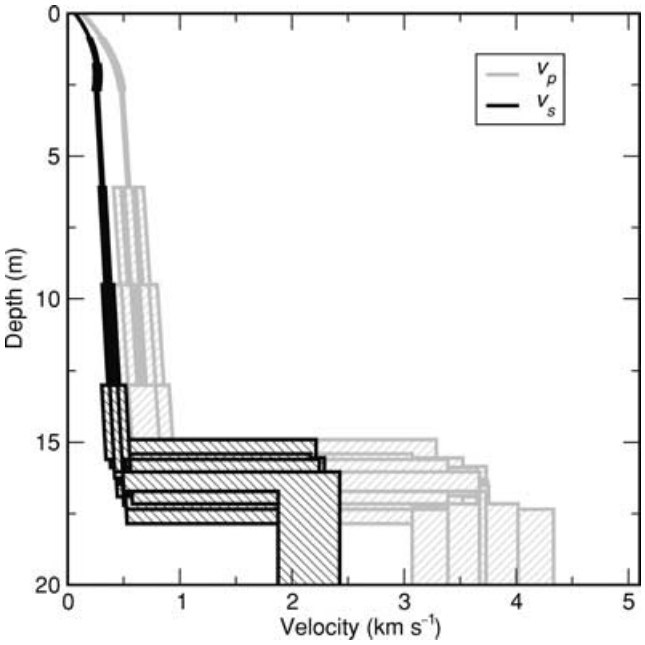

Figure 4. Resolution analysis for data set Bietigheim: The broadness of the hatched areas gives a qualitative measure for the constrainedness of the model parameters, taking trade-off into account. The final model (Fig. 2) was subdivided into 8 sections. The test was applied to the mean of $v_{\rm p}$ and $v_{\rm s}$ in each section and to the depth of the halfspace. The result is shown for a relative increase $\Delta\chi^2/\chi^2(\mathbf{m}_{\rm opt})$ of 2 per cent in the misfit. The figure shows a spurious resolution for $v_{\rm s}$ in the halfspace, for which constraints result from changes in $G_\delta(\omega, p)$ at $p < 1$ s km^{-1}, where the fit is poor.

Table 1. Data set Bietigheim: Drilling log that was taken a few decametres away from the seismic profile (by courtesy of the Wilhelm Fink KG). A driving test at the seismic location could be carried down only to 9 m depth for technical reasons, but confirmed the log for that depth range.

0 m–0.3 m	Topsoil
0.3 m–3 m	Loess and loess loam (Quaternary)
3 m–13 m	Mud and loess loam (Quaternary)
13 m–15 m	Fluvial gravel (Quaternary)
15 m–21 m	Dolomite (Upper Muschelkalk)
Below 21 m	Limestone (Upper Muschelkalk)

3.1.1 How well constrained are the model parameters?

To test for the constrainedness of the subsurface parameters, the model was subdivided into eight independent sections. Fig. 4 shows the results of the resolution analysis for the mean values of v_p and v_s in each section and for the depth of halfspace. The parameter range defined by eq. (21) is given by a hatched area. All tests are superimposed on each other in the plot. The broader the total hatched area is, the less constrained is the corresponding parameter. No damping was applied ($\lambda = 0$ in eq. 13). Data with better fits are given larger weights in the analysis. Weighting all data equally would apparently improve the resolution of v_p at depth greater than 5 m.

S-velocity is well constrained by the (ω, p) coefficients down to 13 m depth. The apparent increase in resolution at depth greater than 2.5 m is due to the increase in thickness of the test sections. The resolution of v_s in the halfspace is spurious. The synthetics that contribute to this are all from $G_\delta(\omega, p)$ with $p < 1\ \mathrm{s\,km^{-1}}$. However, the fit is rather poor in that slowness range. If we remove these data from the test, v_s is almost unconstrained in the halfspace.

From the *P*-wave arrivals we have direct constraints on v_p only down to 5 m and again in the halfspace. Below 5 m v_p is weakly constrained by the (ω, p) coefficients.

Due to the high refractor velocity, the range between 5 m depth and the halfspace is hidden to the first arrivals. They can be explained by a variety of different *P*-velocity models with a refractor anywhere between 15 m and 21 m depth or even shallower, if a low-velocity channel is allowed. The (ω, p) spectrum alone provides no constraint on the depth of halfspace. Only the combination of constraints on v_p from the (ω, p) spectrum and the onset times of the refracted *P*-wave provides a (rather weak) constraint on the halfspace's depth.

Partial derivatives show a significant sensitivity of the fundamental mode to v_p. The main contribution to v_s at shallow depth results from the first higher mode. This possibility is usually neglected in conventional surface-wave inversion.

3.1.2 Synthetic data for the final model

Inferring the field-device response from the data, as defined by eq. (12), allows the calculation of synthetic waveforms for the inversion result. They are compared to the recorded seismograms in Fig. 5 and match them well without explicit waveform fitting.

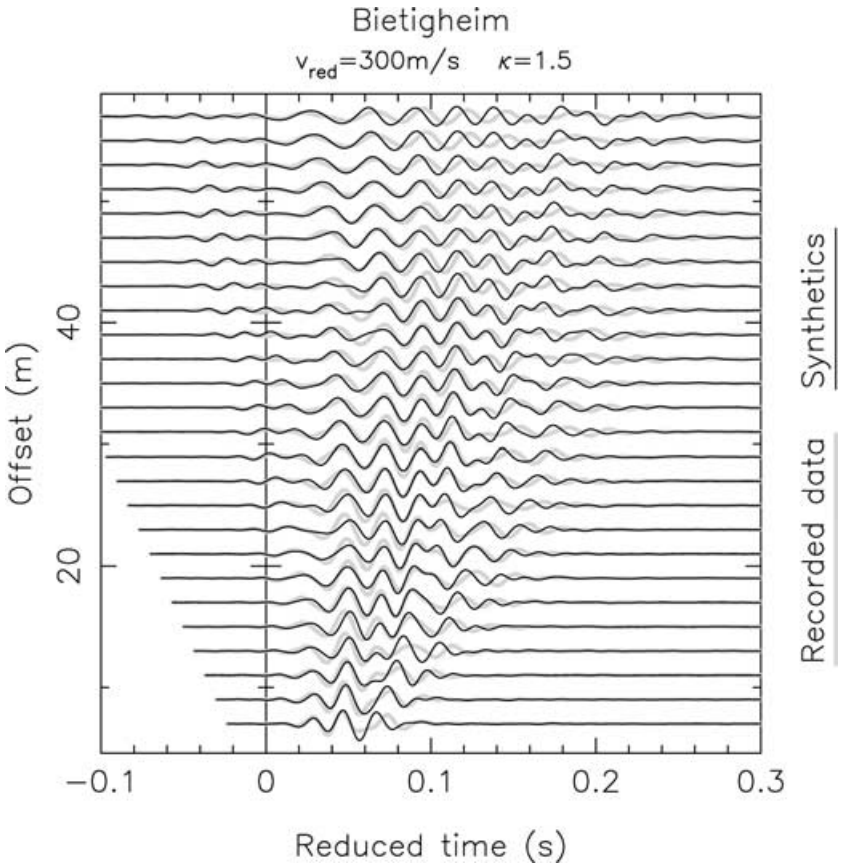

Figure 5. Data set Bietigheim: Synthetic waveforms calculated from the final model (Fig. 2) are superimposed on the recorded seismograms (grey curves). The source wavelet (Fig. 6, top) derived from the (ω, p) spectrum was used for the calculation. The seismograms are given on a reduced timescale with $v_{red} = 300\ \mathrm{m\,s^{-1}}$. Absolute amplitudes are displayed. They are scaled with an offset dependent factor r^κ.

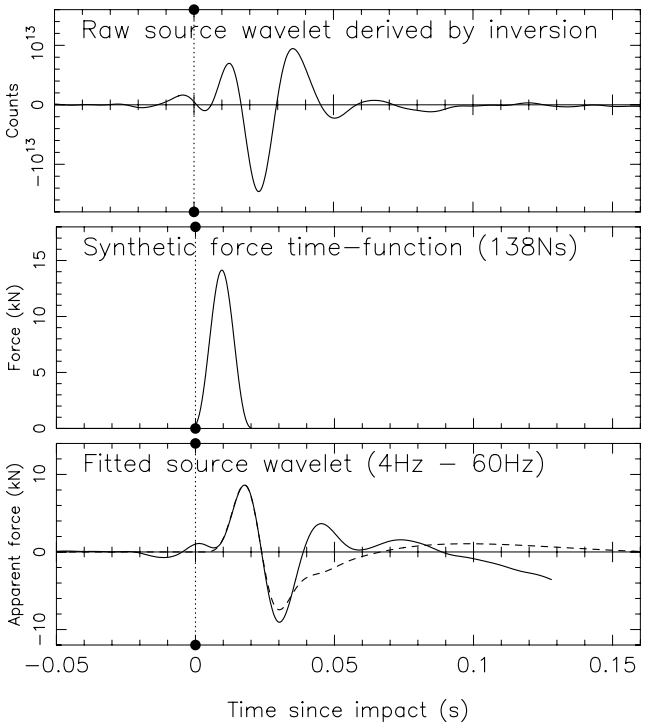

Figure 6. Data set Bietigheim: The top panel shows a source wavelet that was derived from the (ω, p) spectrum by linear regression. The wavelet being deconvolved to a passband from 4 to 60 Hz (due to the filters given in Table 2) is shown in the bottom panel by the solid line. The drift of the signal's baseline for times larger than 0.1 s is due to the long-period noise that is amplified by the deconvolution. The dashed line gives a band-limited version of the force time-function shown in the middle panel with an impulse content of 138 Ns. Times are given relative to the trigger of the seismograph.

The synthetics are slightly shifted if compared with the signal phase of the data, in particular at large offsets. This results from a small departure from the true phase slowness. Waveforms at large offsets are extremely sensitive to phase slowness [much more than the (ω, p) spectrum], because the signal phase is the wavenumber scaled by the offset.

The source wavelet used in the calculation is given in Fig. 6 (top). The signal inferred from the (ω, p) spectrum was scaled by a factor of two in the plot. The linear regression eq. (12) tends to give smaller amplitudes if the fit is not fully satisfactory. A physical interpretation of the source wavelet is given below.

3.2 Data set Berkheim

The second data set was recorded on a hard pitch. The (ω, p) spectrum is shown in Fig. 7. It reveals at least four different Rayleigh modes. The overall characteristics of the wavefield are inverse and anomalous dispersion due to the velocity inversion close to the surface (Paper I).

Fig. 8 shows the model (left) inferred by the inversion of the spectrum and *P*-wave first arrivals (right). The model is remarkably simple. It consists of three structural units. On top is the fast asphalt layer of 20 cm thickness. Below we find a low-velocity section extending to a depth of 5.5 m. There we meet the top of a fast bedrock, which is a jurassic sandstone (Lias α) according to geological background information.

Although the fit might be further improved by allowing for more structural detail in the low-velocity channel, there is no indication of discontinuities. The interface at 90 cm depth was introduced to allow for a gravel foundation, but is not really needed to explain the data. The low-velocity section mainly consists of loam (Bessing, municipality Esslingen, private communication, 1999).

The traveltime curve (Fig. 8, right) is split into arrivals from a fast direct wave in the asphalt layer and waves refracted from deeper regions. Since the fast wave in the asphalt layer radiates refracted waves into the media below, its amplitude decreases rapidly, permitting to read later arrivals. Full space *P*-velocities were used to predict the traveltimes of the direct wave. Since the asphalt thickness is much smaller than the wavelength, this approach is physically questionable. It does, however, not affect the results for the deeper structure.

Fig. 9 shows the spectrum of the Green's function $G_\delta(\omega, p)$ for the final model. It fits the observed data well. The modes are mainly excited in the parts typical for a flexural wave in a thin plate (Angenheister 1950). Since the plate (the asphalt layer) is not free but coupled to the ground below, several overtones contribute to the observed wavefield. Each takes over at an osculation point (Paper I).

3.2.1 How well constrained are the model parameters?

Fig. 10 gives the results of a numerical test of constrainedness. No additional damping is applied ($\lambda = 0$ in eq. 13) and all data are weighted equally. For this analysis the final model was subdivided into nine homogeneous sections and the test is calculated for v_p and v_s in each section and for the asphalt's thickness and the depth of halfspace.

The *S*-velocity in the low-velocity channel is remarkably well constrained by the amplitude maxima in the (ω, p) spectrum. In the opposite, v_s in the asphalt shows a strong trade-off with the thickness of the asphalt layer. The majority of the observable modes represents the flexural wave of the asphalt plate. The phase-velocity of the flexural wave in a thin plate is $v_{ph}(\omega) = \sqrt{\omega K}$. It's dispersion is just sensitive to the constant $K = \sqrt{1/3}\, h\, v_s \sqrt{1-(v_s/v_p)^2}$, which is determined by the seismic velocities and the thickness h of the plate (Brekhovskikh & Goncharov 1985 e.g.).

The depth of the bedrock is well constrained by the frequencies of the osculation points. The halfspace's *S*-velocity is numerically constrained by only a few (ω, p) coefficients around 0.6 s km^{-1} and 12 Hz. They have a remarkably good fit – by pure chance. Omitting these coefficients leaves v_s completely unconstrained in the halfspace.

The *P*-velocity is well constrained for the bedrock only. It is poorly constrained in the asphalt and it may be irrelevant there as noted above. From first arrivals we have some constraint on v_p directly below the asphalt as well as in the halfspace. No waves from depth between 0.5 m and 5.5 m contribute to the first arrivals. v_p is virtually not constrained in that depth range. In this data set the (ω, p) spectrum gives only extremely weak contributions to v_p and these only in the uppermost 50 cm.

3.2.2 Synthetic data for the final model

Fig. 11 shows synthetic waveforms calculated from the final model and a source wavelet. The latter was derived by eq. (12) and is shown in Fig. 12 (top). The synthetic predictions (black) are compared to the recorded seismograms (grey). While the overall waveforms fit

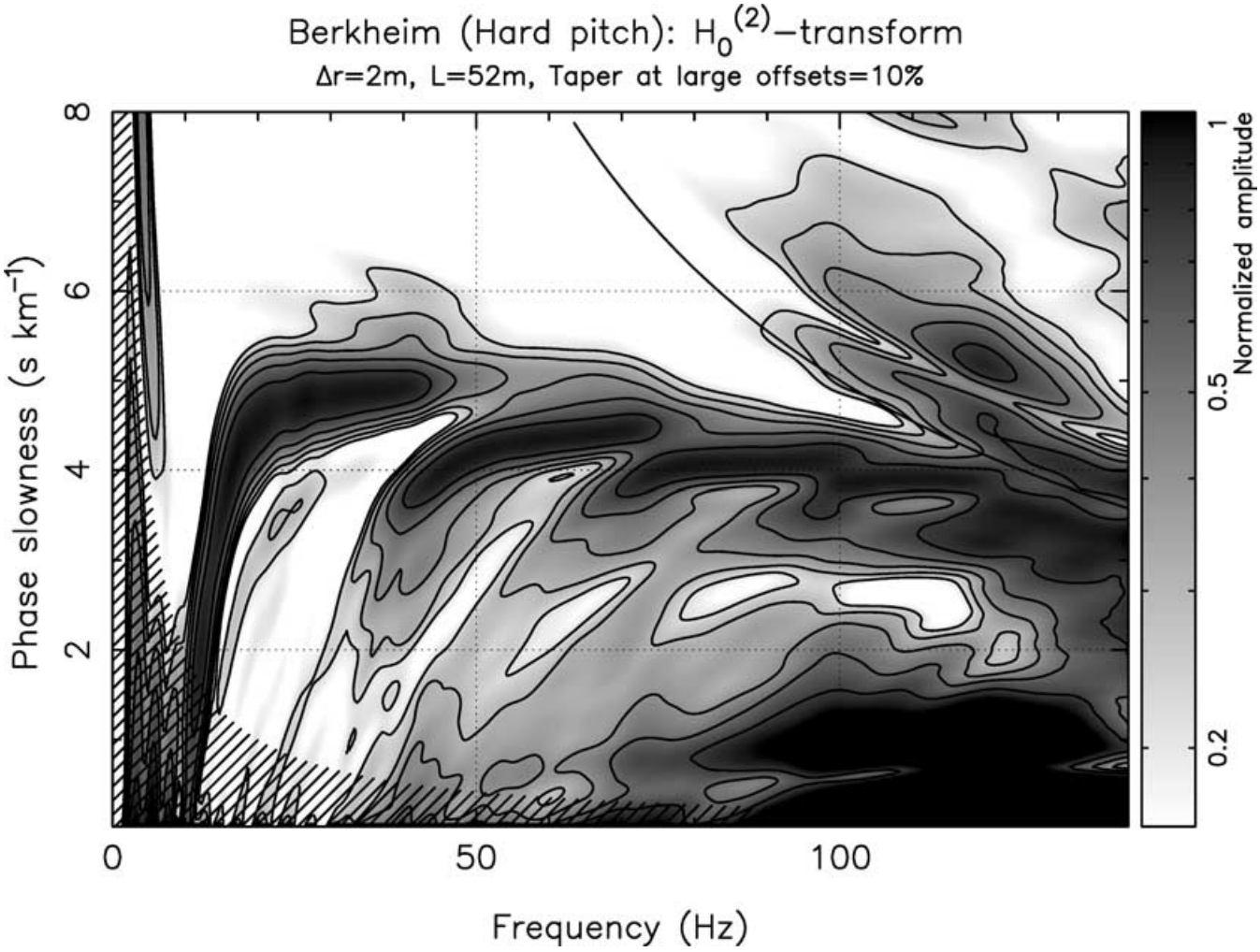

Figure 7. Data set Berkheim: (ω, p) spectrum that was derived from the recorded data by a modified Hankel transformation. We can distinguish at least four different modes in the slowness range from 3 s km^{-1} to 5 s km^{-1}. The signal at 1 s km^{-1} and frequencies larger than 100 Hz may be a *P*-wave from below the asphalt. The hatched area gives a measure of the maximum resolution possible for a spread length of 52 m. The hyperbola in the upper right corner marks the frequency at which aliasing may appear for a given signal phase-slowness due to the largest single-shot geophone-interval of 2 m.

well, traces at large offset indicate that attenuation is still too weak at high frequencies. Q values are between 10 and 50 and should probably be less than 10 in the shallow part of the model.

Fig. 13 (left) shows group-traveltime dispersion-curves for the free modes together with the excitation coefficients for a vertical force at the surface. All modes are excited in group-velocity minima. This explains the spectrogram derived from the recorded data, which is shown beside and was already discussed in Paper I. While the overall wavefield has the character of anomalous and inverse dispersion, the individual free modes have not, in the frequency range where they can be observed. That is obvious from Figs 9 and 13.

3.3 Source time-functions

The field-device response derived by eq. (12) contains not only the geophone's response to ground velocity and instrumental filter responses. It contains also the force time-function of the hammer's impact. The geophone's response can be deconvolved to ground displacement and a natural frequency of 4 Hz. The remaining filter effects (Table 2) lead to a band-limited version of the force wavelet, as shown in the bottom panels of Figs 6 and 12.

The momentum that is transferred by the hammer's impact is always directed into the ground. Thus the true force time-function

Table 2. Filter effects remaining in the deconvolved, but still band-limited version of the force time-function. The parameters apply to a Butterworth characteristic (HP = high-pass, LP = low-pass).

Parameters	Stage
Data set Bietigheim:	
HP 4 Hz, 2 poles	Geophones (after deconvolution)
HP 8 Hz, 2 poles	Data recorder
LP 80 Hz, 4 poles	Signal processing
LP 500 Hz, 4 poles	Data recorder
Data set Berkheim:	
Parameters	Stage
HP 4 Hz, 2 poles	Geophones (after deconvolution)
HP 4 Hz, 2 poles	Data recorder
LP 60 Hz, 4 poles	Signal processing
LP 250 Hz, 4 poles	Data recorder

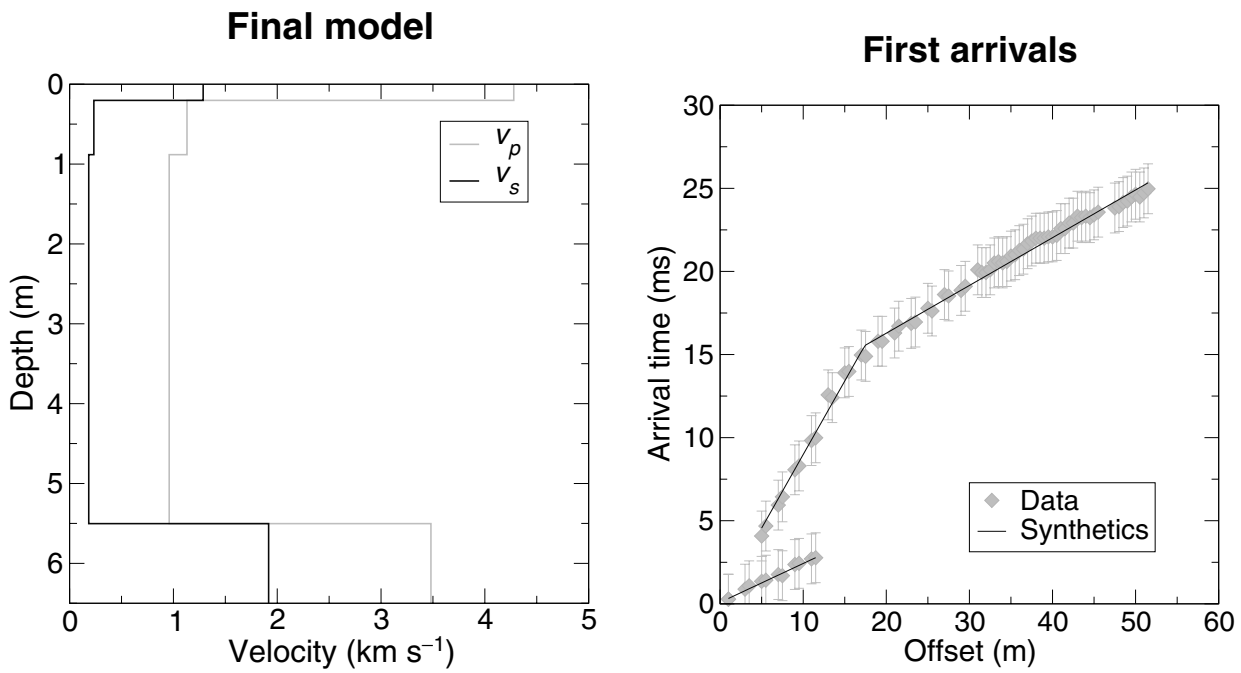

Figure 8. Data set Berkheim: The left graph shows the final model for *P*- and *S*-velocity derived from the (ω, p) spectrum (Fig. 7) and the *P*-wave first arrivals (right graph). The thin layer on top represents the asphalt plate. The right graph shows *P*-wave arrivals (with error bars) picked from the data. They are superimposed by the arrival times predicted by the final model. The traveltimes are subdivided into two groups. The first is due to the direct wave in the asphalt, which dies out at a few metres offset.

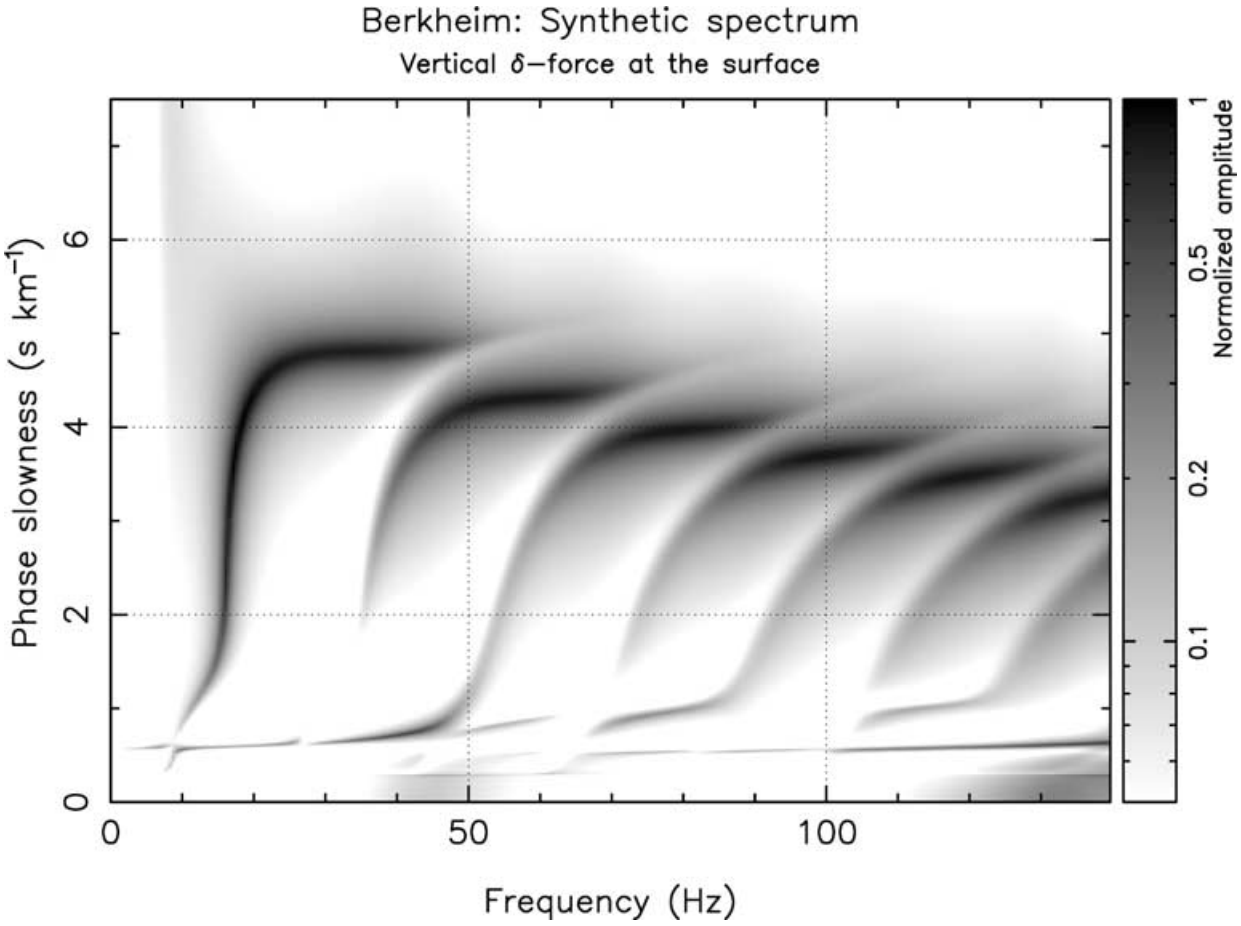

Figure 9. Data set Berkheim: Spectrum of the synthetic Green's function $G_\delta(\omega, p)$ predicted by the final model (Fig. 8). It has the same multi-mode character like the observed wavefield in Fig. 7. Horizontal asymptotes at phase slowness less than 1 s km^{-1} are related to refracted body-wave onsets.

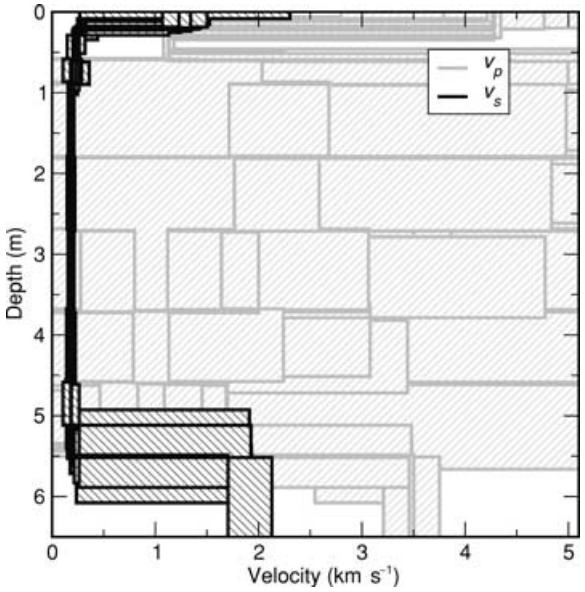

Figure 10. Resolution analysis for data set Berkheim: The broadness of the hatched areas give a qualitative measure for the constrainedness of the model parameters. The broader the area, the less constrained is the parameter, taking trade-off into account. The final model (Fig. 8) was subdivided into 9 sections. The test was applied to the mean of v_p and v_s in each section and for the thickness of the asphalt and the depth of the halfspace. The *P*-velocity in the depth range from 0.5 m to 5.5 m is unconstrained, since there are no first arrivals from that depth range. The result is shown for a relative increase $\Delta\chi^2/\chi^2(\mathbf{m}_{\rm opt})$ of 5 per cent in the misfit. Resolution of v_s in the halfspace is not reliable, since it numerically results only from a few (ω, p) coefficients at 0.6 s km^{-1} and 12 Hz.

is positive. During acceleration of the hammer preceding the actual impact, the force exerted to the ground is reduced. However, this release-force is too small to be observed. The two-sided nature of the shown wavelets is due to the remaining high-pass filters and ringing of the low-pass filters. We can estimate the true force time-function by filtering a reasonable impulse-function with the known characteristics and fitting the result to the wavelet derived from the data. The result of this process is shown in the bottom panels of Fig. 6 for Bietigheim and Fig. 12 for Berkheim. Since the data is high-pass filtered, the transferred momentum cannot be inferred directly. But if we assume that the source time-function is a one-sided impulse, we can obtain at least an estimate. The fit is satisfactory for impulsive wavelets of finite pulse-width and with a momentum of 138 Ns for Bietigheim and 85 Ns for Berkheim. Since the hammer was manually accelerated in addition to gravity and is weakly reflected by the ground, the transferred momentum should be larger than the momentum of an 8 kg weight dropped from 2.5 m height, which is 56 Ns. This condition is satisfied for both fits.

Remarkably the fit is best for input wavelets of finite duration. This suggests a finite duration of the hammer's impact on elastic ground. The impulse on hard asphalt is shorter. The onsets match the trigger time reasonably well.

4 COMPARISON WITH CONVENTIONAL TECHNIQUES

The use of conventional techniques for dispersion analysis gives misleading results when applied to multi-mode data sets (Paper I). For both example data sets the multi-mode character may remain unnoticed. In the case of Bietigheim the first higher mode would be taken for the fundamental mode at frequencies larger than

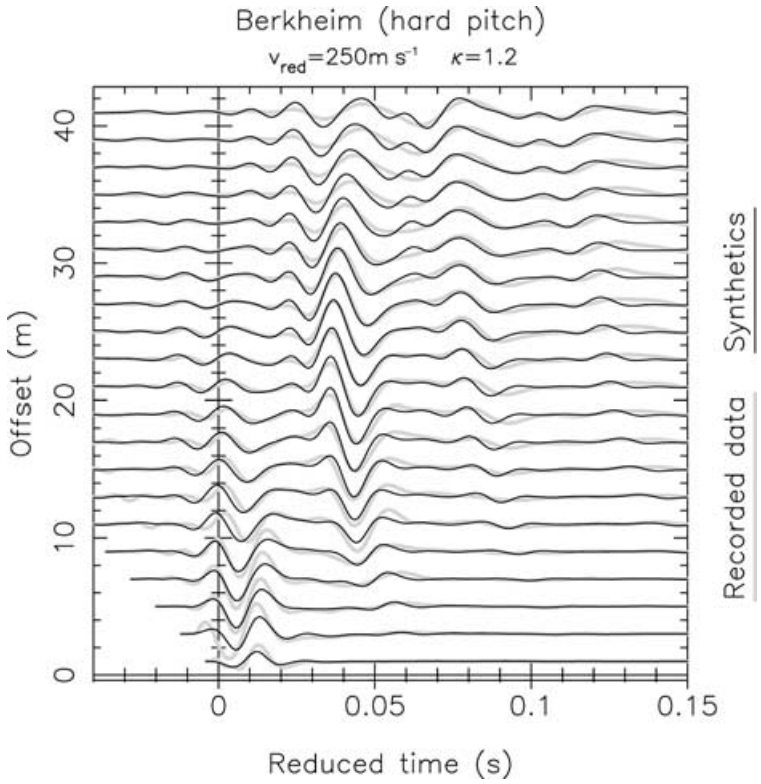

Figure 11. Data set Berkheim: Synthetic waveforms calculated from the final model (Fig. 8) are superimposed on the recorded seismograms (grey curves). The source wavelet (Fig. 12, top) derived from the (ω, p) spectrum was used in the calculation. While the overall waveforms match very well, it is obvious that attenuation of high frequencies is still too weak in the synthetics.

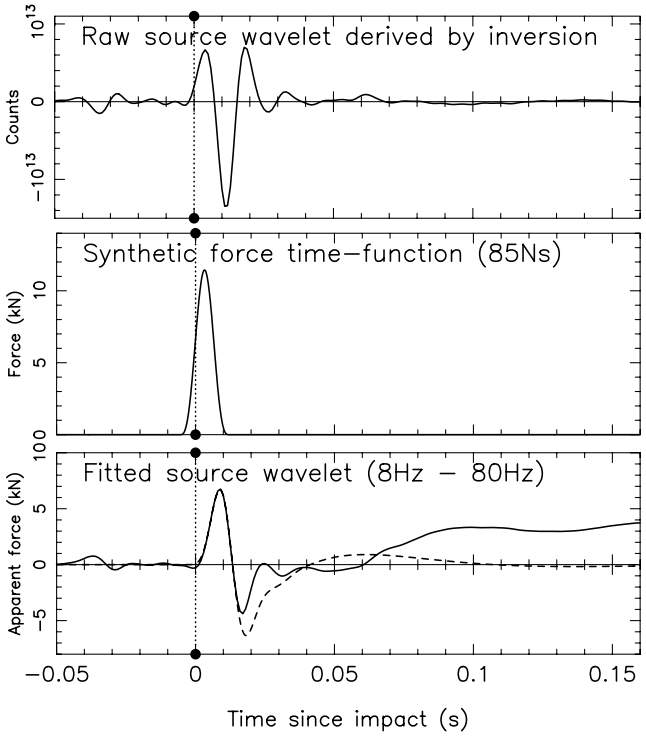

Figure 12. Data set Berkheim: The top panel shows a source wavelet that was derived from the (ω, p) spectrum by linear regression. The wavelet being deconvolved to a passband from 8 to 80 Hz (due to the filters given in Table 2) is shown in the bottom panel by the solid line. The drift of the signal's baseline for times larger than 60 ms is due to the long-period noise that is amplified by the deconvolution. The dashed line gives a band-limited version of the force time-function shown in the middle panel with an impulse content of 85 Ns. Times are given relative to the trigger of the seismograph.

30 Hz. And in the case of Berkheim a fundamental mode with anomalous dispersion would be inferred. I will now investigate the consequences of these misinterpretations with respect to the resulting subsurface model.

4.1 Missing the fundamental mode

For data set Bietigheim a subsurface model was sought that fits the first arrivals and that fits the first higher mode by fundamental-mode dispersion. Fig. 14 shows the (ω, p) spectrum of the data set together with normal-mode dispersion-curves calculated from the model in Fig. 15 (left). The first arrivals are well explained by the v_p model. They give an upper bound on v_p. Regions of lower *P*-velocity could be easily incorporated in the model, because they do not contribute to the first arrivals. On the other hand, the surface wave dispersion gives a lower bound on v_s. By taking the higher mode wrongly as a fundamental mode we overestimate the phase velocity and thus v_s. As a result Poisson's ratio of the final model is too small. Fig. 16 shows that it contains unrealistic negative values. But this is only revealed by the joint inversion. Without fitting the *P*-wave arrivals too, the misinterpretation would remain unnoticed.

4.2 Missing osculation points

In the case of data set Berkheim a model was sought that explains the overall anomalous dispersion by a single fundamental mode and fits the first arrivals. Fig. 17 shows the (ω, p) spectrum of the recorded data together with dispersion curves calculated for the model given in Fig. 15 (right). There is an apparent crossing of dispersion curves at 70 Hz and 4.2 s km^{-1}. However, as Fig. 18 reveals, this is a strong osculation point. The modes's amplitudes for a hammer source are given like error-bars in the figure. The excited half of each of the osculating mode branches contributes to the flexural wave of the asphalt layer. The other half of each mode is related to a channel mode, that has virtually no vertical amplitude at the surface. Thus the false model apparently explains the observed wavefield by a single mode up to 100 Hz.

The position of the transition of fundamental-mode's dispersion to smaller phase slowness with decreasing frequency at approximately 20 Hz, is related to the depth of the high-velocity bedrock. The deeper the bedrock is, the smaller the frequency at which the transition takes place. This is also the case for the higher modes. The regular pattern of modes reaching a plateau near 0.5 s km^{-1} in Fig. 9 is due to this. For this reason it is not possible to find a model that predicts a slow fundamental mode down to 20 Hz and that does

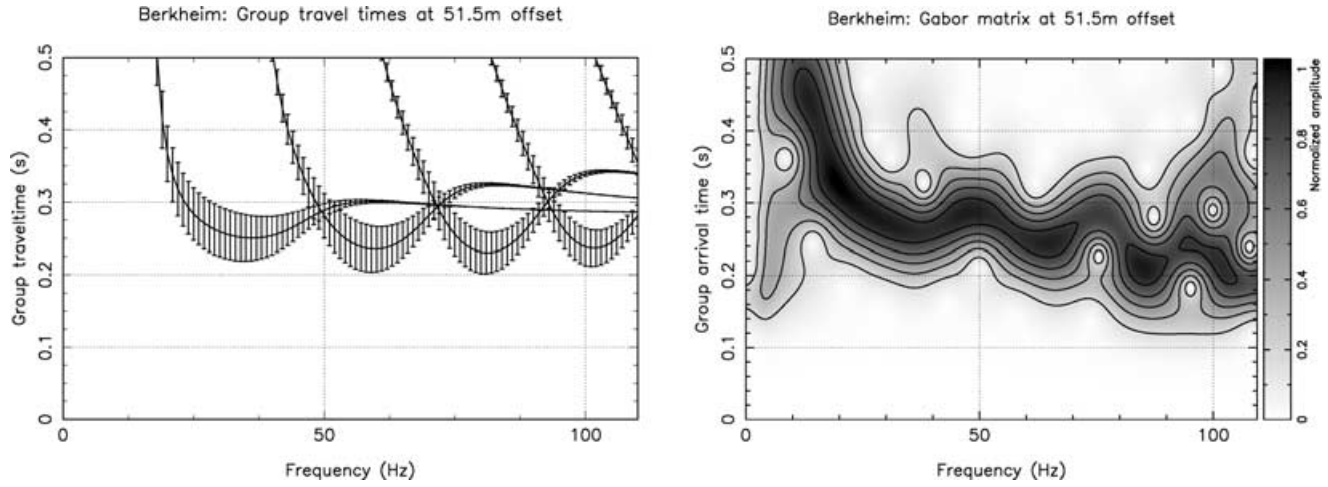

Figure 13. Data set Berkheim: The left graph shows synthetic group-traveltimes for the offset 51.5 m as predicted by the final model (Fig. 8). Excitation amplitudes of the wavefield are calculated for a vertical force and vertical geophones at the surface, and are displayed like error bars. The figure on the right is a spectrogram calculated for the trace at 51.5 m offset. Body waves were removed by tapering. The greyscale gives the signal amplitude. Group arrivals are indicated by dark areas. The observed group traveltimes are well explained by the prediction calculated from the model.

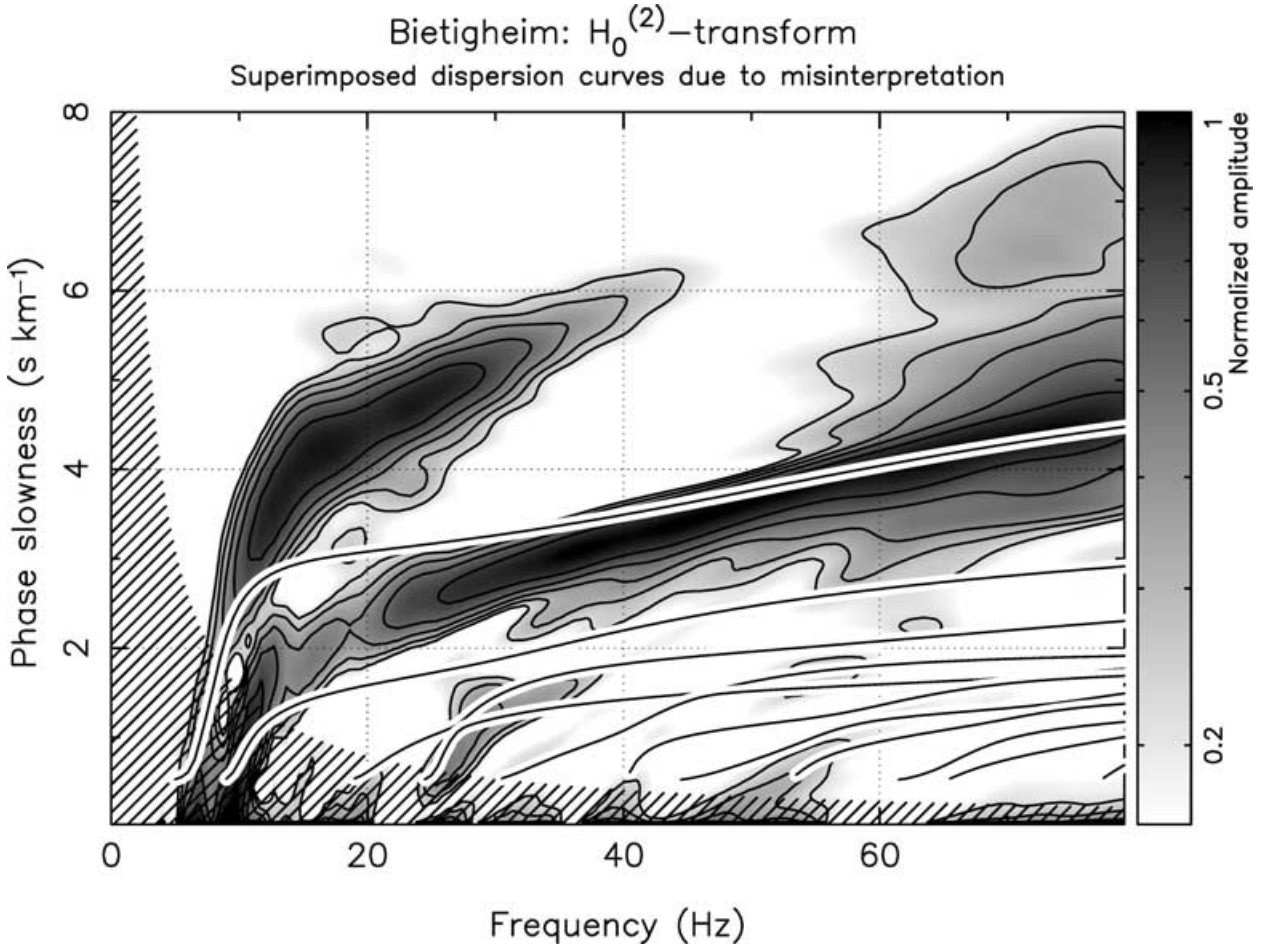

Figure 14. False interpretation of data set Bietigheim: (ω, p) spectrum superimposed by the normal mode dispersion calculated for the model given in Fig. 15 (left). This model predicts a fundamental mode with approximately the dispersion of the dominating first higher mode in the observed data. Additionally it predicts the observed *P*-wave arrival times (not shown).

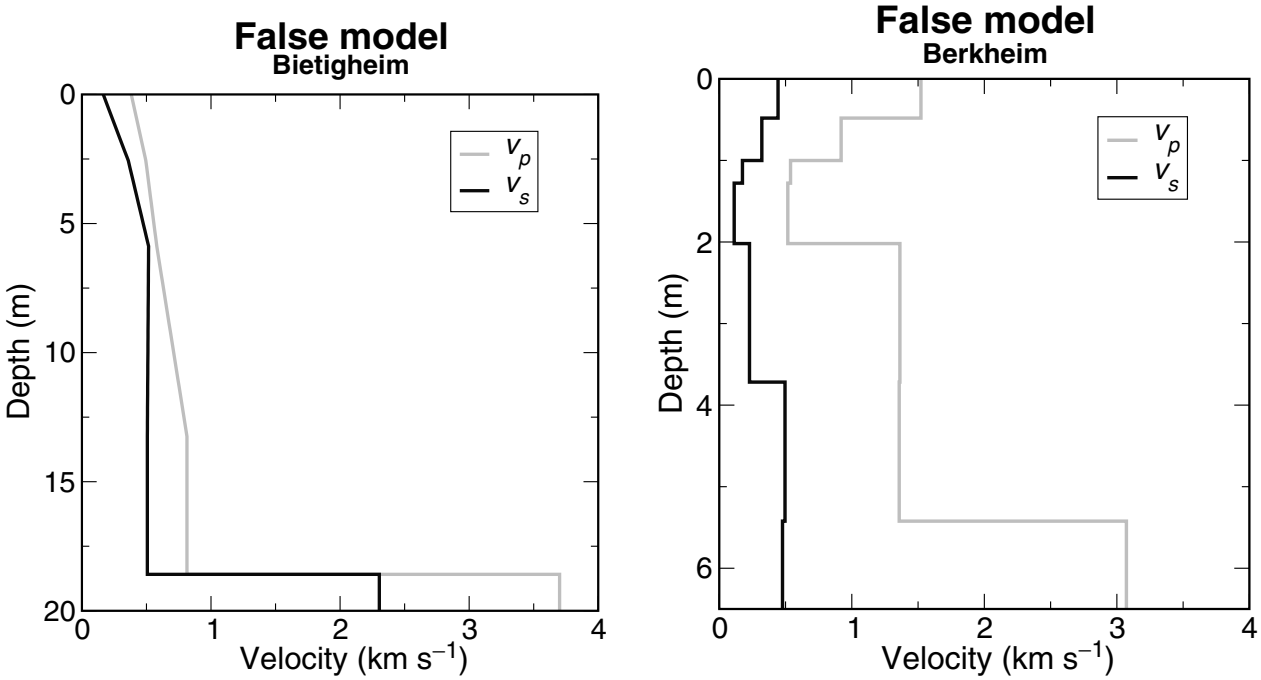

Figure 15. Models resulting from a false interpretation of data sets Bietigheim (left) and Berkheim (right). The normal mode dispersion predicted by these models is given in Figs 14 (Bietigheim) and 17 (Berkheim).

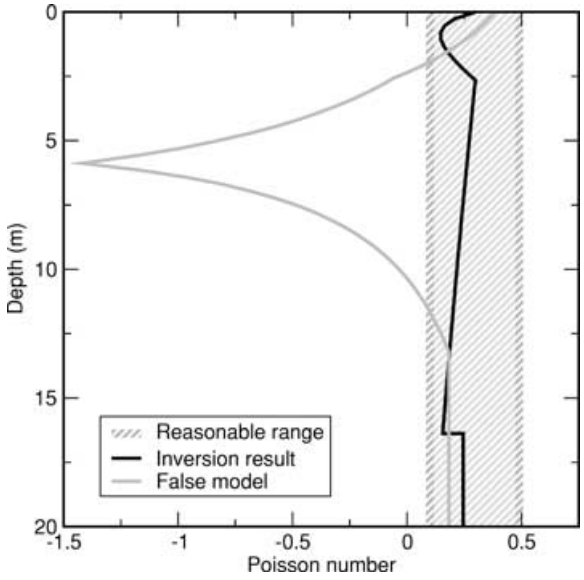

Figure 16. False interpretation of data set Bietigheim: Poisson's ratio for the inversion result (Fig. 2) and the false model (Fig. 15, left). Since the surface-wave dispersion provides a lower bound on v_s and the *P*-wave arrivals provide an upper bound on v_p, the false interpretation (Fig. 14) leads to an unreasonable negative Poisson's ratio. The black curve was calculated for the inversion result shown in Fig. 2. Its detail features may be mainly due to the parametrization of v_p and v_s with polynomials. But the values keep to the range of a reasonable Poisson number.

not have higher modes in the frequency range up to 100 Hz. This must also be the result with conventional inversion and should thus reveal the misinterpretation.

Finally, consider a data set recorded on a site with the properties given in Fig. 15 (right). Due to the amplitude characteristic shown in Fig. 18 it would be impossible to distinguish the fundamental mode from the higher mode even in the (ω, p) spectrum. Fitting dispersion curves alone would be almost impossible in that case.

The normal-mode synthetics presented in this section are calculated with a strinkingly robust code (FLSPHER by Wolfgang Friederich). It makes use of second order minors for the the integration of the elastodynamic system of ordinary differential equations (Friederich & Dalkolmo 1995) in combination with a root-count technique (Woodhouse 1988). Both features are essential to obtain accurate results in the vicinity of osculation points.

5 CONCLUSIONS AND OUTLOOK

The two data sets for which I show inversion results are examples for typical classes of subsurface structure out of a larger number of field data. One reveals a smooth non-linear increase of seismic velocity with depth. The other includes a pronounced low-velocity channel under a fast lid. Both reveal a large variation of seismic velocities from ground level to bedrock. Constraints on v_s are remarkably good down to bedrock in both cases. The depth of bedrock is reasonably constrained, which would be impossible from *P*-wave refraction alone.

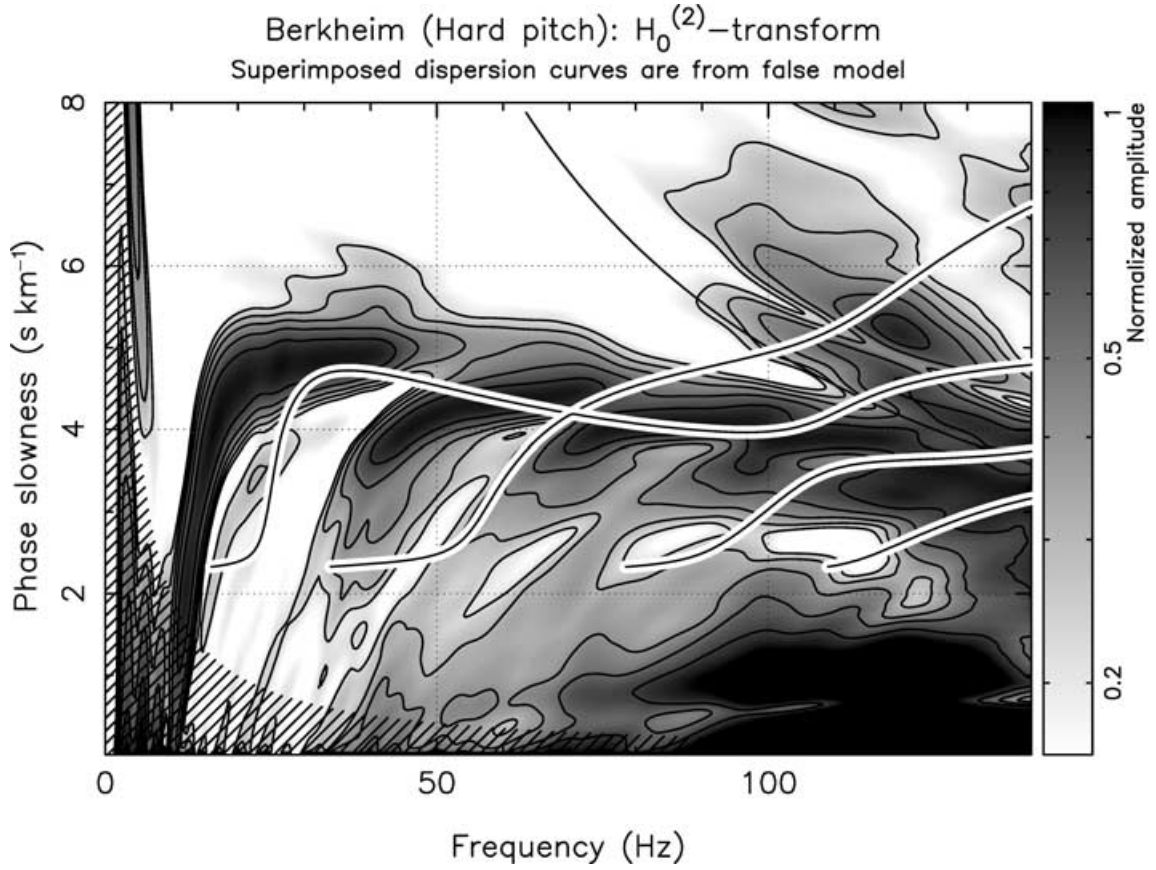

Figure 17. False interpretation of data set Berkheim: (ω, p) spectrum superimposed by the normal mode dispersion of of the model given in Fig. 15 (right). Notice, that the osculation point at 70.3 Hz and 4.19 s km^{-1} cannot be recognized in the excited wavefield (Fig. 18). The fundamental mode above 70 Hz and the first higher mode below 70 Hz are not excited by a hammer source at the surface. With respect to the excited wavefield, the signal observed at Berkheim is explained by the false model apparently with a single mode up to 100 Hz.

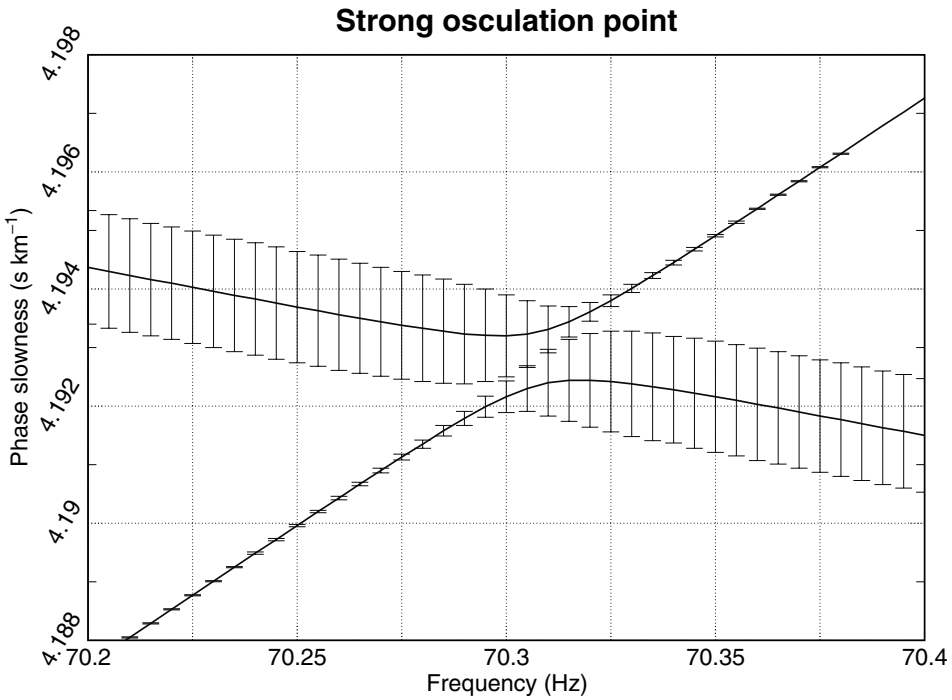

Figure 18. False interpretation of data set Berkheim: The false model given in Fig. 15 (right) produces a strong osculation point at 70.3 Hz and 4.19 s km^{-1}. The curves give the normal mode dispersion and the amplitude excitation for a vertical force and vertical geophones at the surface is displayed like error-bars. The apparent crossing of normal-mode dispersion-curves in Fig. 17 is an osculation point. Notice, that with respect to the excited wavefield, the observed signal is explained by the false model apparently with a single fundamental mode up to 100 Hz.

The examples prove that the proposed inversion scheme is a robust means to retrieve subsurface velocities from Rayleigh-wave data even without *a priori* information. In particular, it performs well with multi-mode wavefields that are often observed in the shallow domain. That is because it does not require to identify normal-modes prior to inversion, even though this is possible in the (ω, p) domain for many cases. Rather than fitting dispersion curves, we search for a subsurface model that predicts large amplitudes at the correct locations in the (ω, p) transform for a given source and ground-displacement component. Thus the full wavefield-content including the dispersion of higher modes and their true amplitudes is exploited. Since the inversion needs not evaluate the Fourier–Bessel expansion, as would be necessary for waveform fitting, this method can be numerically more efficient by up to a factor of ten.

In comparison, conventional dispersion-curve fitting is problematic when applied to the example data sets. In a careful study that exploits *P*-wave arrivals too, systematic errors might be noticeable. But both data sets are still likely to give misleading results with conventional methods.

The subsurface models resulting from the proposed method predict the recorded waveforms surprisingly well, although no waveform fitting was applied. This can serve as an additional indicator of quality. Source wavelets inferred from the data without constraining them to a certain curve form are physically reasonable. The force time-function is well fitted by a one-sided pulse of finite duration. This indicates a parametrization that will be helpful in waveform fitting.

Apart from direct interpretation, the final models obtained by this method can serve as initial models for a subsequent waveform inversion. Waveform fitting can provide even stronger constraints on velocity because the signal phase is an integral of phase slowness over offset. Additionally it can provide more useful constraints on attenuation. Attenuation in final models from the technique proposed here is typically still too weak due to conservative assumptions in the initial models. Waveforms indicate $Q < 10$. But waveform fitting is not possible without a method to find an appropriate initial model. With the proposed technique this final step in seismogram interpretation should now be within reach.

ACKNOWLEDGMENTS

I gratefully thank Erhard Wielandt and Wolfgang Friederich for inspiring discussions and constructive comments on the manuscript. I am grateful to Stefan Hecht and Gunther Reimann who contributed in the field work and to Wolfgang Friederich who provided his code FLSPHER for the computation of normal modes. I thank Wolfgang Rabbel and an anonymous reviewer for their encouraging comments. This study was supported by the Institute of Geophysics at the University of Stuttgart.

REFERENCES

Angenheister, G., 1950. Fortschreitende elastische Wellen in planparallelen Platten, *Gerl. Beitr. z. Geophys.,* **61,** 296–308.

Bachrach, R., Dvorkin, J. & Nur, A., 1998. High-resolution shallow-seismic experiments in sand, Part II: Velocities in shallow unconsolidated sand, *Geophysics,* **63,** 1234–1240.

Bachrach, R., Dvorkin, J. & Nur, A.M., 2000. Seismic velocities and Poisson's ratio of shallow unconsolidated sands, *Geophysics,* **65,** 559–564.

Brekhovskikh, L. & Goncharov, V., 1985. *Mechanics of Continua and Wave Dynamics,* Springer, Berlin.

Burkhardt, H., Mörig, R. & Schütt, R., 1992. Laboratory investigations on rock samples to establish further fundamentals for the lithological interpretation of seismic measurements, in *Absorption of seismic waves,*

Vol. 397, pp. 243–279, German Society for Petroleum and Coal Chemistry.

Dziewonski, A.M. & Anderson, D.L., 1981. Preliminary reference earth model, *Phys. Earth planet. Inter.,* **25,** 297–356.

Forbriger, T., 2001. *Inversion flachseismischer Wellenfeldspektren,* Dissertation, Institut für Geophysik, Universität Stuttgart, http://elib.uni-stuttgart.de/opus/volltexte/2001/861.

Forbriger, T., 2003. Inversion of shallow-seismic wavefields: I. Wavefield transformation, *Geophys. J. Int.,* **153,** 720–735 (this issue).

Friederich, W. & Dalkolmo, J., 1995. Complete synthetic seismograms for a spherically symmetric earth by a numerical computation of the Green's function in the frequency domain, *Geophys. J. Int.,* **122,** 537–550.

Gucunski, N. & Woods, R.D., 1991. Inversion of Rayleigh wave dispersion curve for SASW test, in, *Soil Dynamics and Earthquake Engineering V,* pp. 127–138, Institut für Bodenmechanik und Felsmechanik der Universität Karlsruhe, Elsevier Applied Science, London, New York.

Jackson, D.D., 1976. Most squares inversion, *J. geophys. Res.,* **81,** 1027–1030.

Jones, R., 1958. In-situ measurement of the dynamic properties of soil by vibration methods, *Geotechnique,* **8,** 1–21.

Matthews, M., Hope, V. & Clayton, C., 1996. The use of surface waves in the determination of ground stiffness profiles, *Proc. Instn Civ. Engrs Geotech. Engng,* **119,** 84–95.

Nazarian, S., 1984. In Situ Determination of Elastic Moduli of Soil Deposits and Pavement Systems by Spectral-Analysis-of-Surface-Waves Method, *PhD thesis*, The University of Texas, Austin.

Roth, M. & Holliger, K., 2000. The non-geometric $\overline{P}S$ wave in high-resolution seismic data: observations and modelling, *Geophys. J. Int.,* **140,** F5–F11.

Schön, J.H., 1998. *Physical properties of rocks: Fundamentals and principles of petrophysics,* Vol. 18, of Handbook of Geophysical Exploration, Elsevier Science Ltd., Oxford, 2nd edn.

Spencer, J.W., Jr, 1981. Stress relaxations at low frequencies in fluid-saturated rocks: Attenuation and modulus dispersion, *J. geophys. Res.,* **86,** 1803–1812.

Stokoe, K.H., II, & Nazarian, S., 1983. Effectivness of ground improvement from spectral analysis of surface waves, in, *Proceedings of the Eigth European Conference on Soil Mechanics and Foundation Engineering,* Helsinki.

Stümpel, H., Kähler, S., Meissner, R. & Milkereit, B., 1984. The use of seismic shear waves and compressional waves for lithological problems of shallow sediments, *Geophysical Prospecting,* **32,** 662–675.

Telford, W.M., Geldart, L.P. & Sheriff, R.E., 1990. *Applied Geophysics,* Cambridge University Press, 2nd edn.

Tokimatsu, K., Tamura, S. & Kojima, H., 1992. Effects of multiple modes on Rayleigh wave dispersion characteristics, *J. Geotech. Engng.,* **118,** 1529–1543.

Ungerer, J., 1990. *Berechnung von Nahfeldseismogrammen mit der Reflektivitätsmethode,* Diplomarbeit, Institut für Geophysik, Universität Stuttgart, Germany.

Woodhouse, J., 1988. The calculation of eigenfrequencies and eigenfunctions of the free oscillations of the earth and the sun, in *Seismological Algorithms,* pp. 321–370, ed. Doornbos, D.J., chap. IV.2, Academic Press, London.