JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 93, NO. B10, PAGES 12,055–12,065, OCTOBER 10, 1988

Large-Scale Waveform Inversions of Surface Waves for Lateral Heterogeneity 1. Theory and Numerical Examples

ROEL SNIEDER

Department of Theoretical Geophysics, University of Utrecht, The Netherlands

Surface wave scattering theory is presented as a new method for analyzing teleseismic surface wave data. Using surface wave scattering integrals the effect of lateral heterogeneity both on the surface wave coda generation and on the direct surface wave is described. Since the employed scattering theory for the forward problem is linear, the inverse problem can conveniently be solved in the least squares sense using an iterative matrix solver. For waveform inversions of the direct surface wave, only near forward scattering contributes. For this case the isotropic approximation is introduced, which makes it possible to retrieve phase velocity information from scattering theory. It is shown that for practical waveform inversions the resulting system of linear equations is extremely large and how row action methods can be used conveniently for carrying out the inversion on moderate size computers. The performance of the inversions is illustrated with two numerical examples. In the first example the surface wave coda generated by one point scatterer is inverted. It is shown that the reconstruction in this case is similar to Kirchhoff migration methods as used in exploration seismics. In the second example, ray geometrical effects (focusing and phase shifting) are obtained from the linear inversion with scattering theory. It follows from this example that linear waveform inversion can simultaneously fit the amplitude and the phase of surface wave data.

1. INTRODUCTION

Standard surface wave analysis proceeds by extracting path-averaged group or phase velocities from surface wave data using dispersion analysis. If sufficient data are available, these path-averaged dispersion data can be used to determine the local phase or group velocity. Mathematically, this approach relies on the great circle theorem [*Backus*, 1964; *Jordan*, 1978; *Dahlen*, 1979] or more accurately on the minor arc theorem [*Romanowicz*, 1987]. These theorems state that surface waves are only influenced by the integral of the phase or group velocity over the source receiver great circle (or minor arc). This is justified if the lateral heterogeneity is smooth on a scale of a wavelength of the surface waves under consideration.

In practice, this condition may not be satisfied. For example, a 30-s Rayleigh wave has a wavelength of approximately 120 km. In continents the lateral variation on this scale can be considerable, so that the use of the great circle (minor arc) theorem and the related dispersion measurements are not justified. Surprisingly, this well-known fact is widely ignored, and in some cases, dispersion analysis is used over structures which have the same length scale as the surface waves [e.g., *Panza et al.*, 1980; *Calcagnile and Scarpa*, 1985]. If the structure is not smooth on a scale of a wavelength, surface wave scattering and multipathing may occur. This is documented for reflection of surface waves at a continental margin by *Levshin and Berteussen* [1979] and *Bungum and Capon* [1974]. Linearized scattering theory can be used to describe these effects. This theory is developed both for a flat geometry [*Snieder*, 1986*a*,*b*], and for a spherical geometry [*Snieder and Nolet*, 1987].

Scattered surface waves must to some degree be responsible for the generation of the surface wave coda, and it would be fruitful to extract this information from the surface wave coda. *Snieder* [1986a] presents a holographic inversion scheme for the surface wave coda, reminiscent of migration procedures in exploration seismics. This inversion method has been applied successfully to image the surface wave reflections from a concrete dam on a tidal flat [*Snieder*, 1987*a*]. In order to achieve this, several severe approximations have been used, and it is desirable to give waveform inversion for surface data a firmer theoretical basis. This paper serves to provide a rigorous waveform-fitting method for surface waves, based on surface wave scattering theory. This inversion is set up as a huge matrix problem, and it is shown how solutions can be found iteratively.

There is, however, more to be gained from surface wave scattering theory than an analysis of the surface wave coda. Surface wave scattering theory can also be used to describe the distortion of the direct wave due to lateral heterogeneity [*Snieder*, 1987*b*]. This allows not only for accurate forward modeling of the direct surface wave in the presence of lateral heterogeneity but also for a waveform inversion of the direct surface wave train. In this way both amplitude and phase information can be used.

A waveform inversion of surface wave data was first attempted by *Lerner-Lam and Jordan* [1983], who linearly fitted higher-mode surface waves with a laterally homogeneous model. *Nolet et al.* [1986*a*] extended this method to incorporate nonlinear effects and lateral inhomogeneity. However, they only used the phase information of the surface waves. *Yomogida and Aki* [1987] used the Rytov field to fit both the amplitude and phase of fundamental mode Rayleigh wave data. The starting point of *Yomogida and Aki* [1987] is the two-dimensional wave equation. One can argue that their method lacks rigor because it is not clear that surface waves satisfy the two-dimensional wave equation. *Tanimoto* [1987] determined a global model for the S velocity in terms of spherical harmonics up to order 8 using long-period higher-mode waveforms. In computing the synthetics he used the great circle theorem to compute the phase shift, and he ignored focusing effects. Because of the low order of his spectral expansion ($l \leq 8$), ray theory could be used for this inversion. This means that up to this point, all waveform inversions for surface waves relied either on ray theory or on the two-dimensional wave equation.

In this paper it is shown how linear scattering theory can be used for waveform fitting of the direct surface wave by the reconstruction of a two-dimensional phase velocity field. The derivation uses the full equations of elasticity, and neither uses ray

Paper number 7B6093.
0148-0227/88/007B-6093$05.00

theory nor the two-dimensional wave equation. Specifically, there is no need to assume any smoothness properties of the medium. In fact, in section 7 a numerical example is shown of the distortion of the direct surface wave by a structure with sharp edges. A restriction of this inversion method is that small scattering angles are assumed. This can in practice be realized by time windowing the data.

In section 2 some elements of surface wave scattering theory are revisited. The isotropic approximation, which allows the determination of phase velocities from scattering theory, is introduced in section 3. Section 4 features a method to invert the resulting scattering integral. Due to the extremely large size of the resulting matrix equation this is not without problems, and in section 5, several tricks are shown to make these computations feasible on systems as small as a super minicomputer. Unfortunately, the surface wave inversion problem is in reality nonlinear, and the assumption of linearity is only justified for reference models which are sufficiently close to the real Earth. It is therefore advantageous to perform a nonlinear inversion (using ray theory) first (section 6) in order to find a smooth reference model for the subsequent linear inversion. (In this paper this linear inversion is referred to as "Born inversion.") In section 7 it is shown that a more or less realistic distribution of scatterers produces a realistic looking coda but also that sharp lateral heterogeneity may severely distort the direct surface wave. Examples of inversions for a point scatterer and for ray geometrical effects (phase shifting and focusing) are presented in the last two sections. Application of this technique to surface wave data recorded with the Network of Autonomously Recording Seismographs (NARS) are presented by *Snieder*, [this issue] (hereafter referred to as paper 2).

Throughout this paper the limitations of surface wave scattering are assumed [*Snieder and Nolet*, 1987]; that is, it is assumed that the heterogeneity is weak and that the far-field limit can be used. In order to transcend these limitations a considerable amount of theoretical work remains to be done. For reasons of simplicity, only vertical component fundamental mode data are assumed, but this restriction is not crucial. Note that this does not mean that the fundamental Love wave need not be considered, because in general a double-couple source excites Love waves, which may be converted by the heterogeneity to Rayleigh waves.

2. SURFACE WAVE SCATTERING THEORY

A dyadic decomposition of the surface wave Green's function [*Snieder*, 1986*a*; *Snieder and Nolet*, 1987] has allowed compact expressions for both the direct and the scattered surface waves. In this section, elements of surface wave scattering theory are briefly presented. Throughout this paper a spherical geometry is assumed, and computations are performed to leading order of ka, where k is the wave number and a the circumference of the Earth. As shown by *Snieder and Nolet* [1987], the unperturbed surface wave excited by a moment tensor $\mathbf{M}$ can be written as a sum over surface wave modes (with index ν):

$$\mathbf{u}^0(r,\theta,\phi) = \sum_\nu \mathbf{p}^\nu(r,\mu_R)\,\frac{e^{i(k_\nu a\,\Delta+\frac{\pi}{4})}}{\sqrt{\sin\Delta}}\,(\mathbf{E}^{\nu*}(r_S,\mu_S):\mathbf{M}) \qquad (1)$$

In this expression, μ_S and μ_R are the azimuths of the source receiver minor arc at the source and receiver, respectively, counted anticlockwise from south, while Δ is the epicentral distance. In this paper we shall only be concerned with the fundamental modes, so that the (Greek) mode indices are usually omitted. The polarization vector $\mathbf{p}$ is for Love waves given by

$$\mathbf{p}_L = -(l+\tfrac{1}{2})\,W(r)\,\hat{\boldsymbol{\phi}} \qquad (2a)$$

and for Rayleigh waves by

$$\mathbf{p}_R = (l+\tfrac{1}{2})\,V(r)\,\hat{\boldsymbol{\Delta}} - iU(r)\hat{\mathbf{r}} \qquad (2b)$$

where $\hat{\mathbf{r}}$, $\hat{\boldsymbol{\Delta}}$, and $\hat{\boldsymbol{\phi}}$ are unit vectors in the vertical, radial, and transverse direction, respectively. The eigenfunctions U, V, and W of the Earth's normal modes are defined by *Gilbert and Dziewonski* [1975]. The eigenfunctions are assumed to be normalized as by *Snieder and Nolet* [1987]:

$$\tfrac{1}{2}\int \rho(r)\left[U^2(r)+l(l+1)V^2(r)\right]r^2dr =$$
$$\tfrac{1}{2}\int \rho(r)\,l(l+1)W^2(r)\,r^2dr = \left[\frac{l+\frac{1}{2}}{2\pi}\right]^{\frac{1}{2}} /4\omega u_g \qquad (3)$$

The angular quantum number l is related to the wave number by the relation $ka=l+\frac{1}{2}$, and u_g is the angular group velocity of the mode under consideration. The excitation tensor $\mathbf{E}$ in (1) can be expressed in the polarization vector at the source

$$\mathbf{E}(r_S,\mu_S) = \left[\hat{\mathbf{r}}\partial_r + i\frac{(l+\frac{1}{2})}{r}\hat{\boldsymbol{\Delta}}_S\right]\mathbf{p}(r_S,\mu_S) \qquad (4)$$

The perturbation of the wave field due to the lateral heterogeneity can be expressed as a double sum over incoming (σ) and scattered (ν) surface waves modes [*Snieder and Nolet*, 1987]

$$\mathbf{u}^1(r,\theta,\phi) = \sum_{\nu,\sigma}\iint \mathbf{p}^\nu(r,\mu_R')\,\frac{e^{i(k_\nu a\,\Delta_2+\frac{\pi}{4})}}{(\sin\Delta_2)^{\frac{1}{2}}}\,\tilde{V}^{\nu\sigma}(\theta',\phi')\times$$
$$\times\,\frac{e^{i(k_\sigma a\,\Delta_1+\frac{\pi}{4})}}{(\sin\Delta_1)^{\frac{1}{2}}}\,(\mathbf{E}^{\sigma*}(r_S,\mu_S').\mathbf{M})\,d\Omega' \qquad (5)$$

The surface wave distortion is expressed as a scattering integral over the horizontal extend of the heterogeneity (θ',ϕ'). The minor arc from the source to the heterogeneity (θ',ϕ') defines the azimuth μ_S' at the source and the angular distance Δ_1, while the minor arc from (θ',ϕ') to the receiver defines the azimuth μ_R' at the receiver and the angular distance Δ_2. The interaction matrix $\tilde{V}^{\nu\sigma}$ describes the coupling between the modes ν and σ. For isotropic perturbations in the density $\delta\rho$ and the Lamé parameters $\delta\lambda$ and $\delta\mu$ the interaction matrix depends only on frequency and the scattering angle Ψ defined by

$$\Psi = \mu_{out} - \mu_{in} \qquad (6)$$

(μ_{out} and μ_{in} are the azimuths of the incoming and scattered wave at the scatterer.) Extensions for perturbations of interfaces and gravitational effects are given by *Snieder and Romanowicz* [1988], while the effects of anisotropy are discussed by *Romanowicz and Snieder* [1988].

For perturbations in the density and the Lamé parameters the interaction terms are depth integrals containing the heterogeneity and the modes under consideration [*Snieder and Nolet*, 1987]. For example, the Love wave to Rayleigh wave conversion ($R \leftarrow L$) is given by

$$\tilde{V}_{RL} = (l_R+\tfrac{1}{2})(l_L+\tfrac{1}{2})\int\Big[-V_R W_L\,\delta\rho\omega^2$$
$$+\left(\frac{1}{r}U_R+\partial_r V_R\right)(\partial_r W_L)\delta\mu\Big]\,r^2dr\,\sin\psi$$

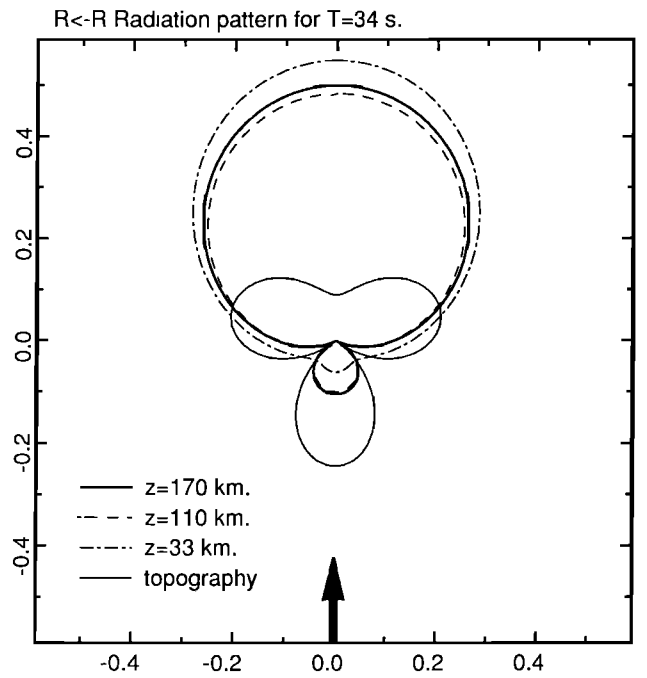

Fig. 1*a*. Radiation pattern for the interaction of the fundamental Rayleigh mode with itself for surface topography (3 km) and for S velocity perturbations extending down to 170 km ($\delta\beta/\beta$=4%), 110 km ($\delta\beta/\beta$=4%), and 33 km ($\delta\beta/\beta$=12%). The effective size of the scatterer is 100×100 km^2. The direction of the incoming wave is shown by an arrow.

$$+ (l_R+\tfrac{1}{2})^2(l_L+\tfrac{1}{2})^2 \int V_R W_L \delta\mu \, dr \, \sin 2\psi \qquad (7)$$

For Love-Love wave interactions (LL) or Rayleigh-Rayleigh wave interactions (RR) a similar dependence on the scattering angle exists

$$\tilde{V}_{RR \text{ or } LL} = \tilde{V}^{(0)}_{RR} + \tilde{V}^{(1)}_{RR \text{ or } LL} \cos\Psi + \tilde{V}^{(2)}_{RR \text{ or } LL} \cos 2\Psi \qquad (8)$$

In Figures 1*a* and 1*b* the radiation patterns are shown for interactions of the fundamental Rayleigh wave with itself and for conversion from the fundamental Love wave to the fundamental Rayleigh wave. These radiation patterns are shown for surface topography [*Snieder*, 1986*b*; *Snieder and Romanowicz*, 1988], and for a constant relative perturbation in the S velocity $\delta\beta/\beta$ down to different depths with an unperturbed density ($\delta\rho$=0). The perturbations in the Lamé parameters are equal.

It is shown by *Snieder* [1986*b*] that the interaction terms for forward scattering and unconverted waves are proportional to the perturbation of the phase velocity δc. Using the normalization (3), equation (9.3) of *Snieder* [1986*b*] can be written as

$$\frac{\delta c}{c} = \left[\frac{2\pi}{l+\tfrac{1}{2}}\right]^{\frac{1}{2}} \frac{1}{(l+\tfrac{1}{2})} \tilde{V}^{\text{unconverted}}(\Psi=0) \; . \qquad (9)$$

Up to this point, it has been assumed that the real Earth can be treated as a radially symmetric reference model (producing a seismogram $\mathbf{u}^0$), with superposed lateral inhomogeneities (leading to the seismogram distortion $\mathbf{u}^1$). However, as shown in *Snieder* [1986*a*], the theory can also be formulated for a smoothly varying reference model, with embedded heterogeneities. (Smooth means that the lateral variation is small on a scale of one horizontal wavelength.) In that case the phase terms and the geometrical spreading terms of the propagators follow from ray theory [*Snieder*, 1986*a*]. Solving the ray tracing equations is a cumbersome affair, and as long as the inhomogeneity of the reference medium is sufficiently weak, the ray geometrical effects can be expressed as simple line integrals over the minor arc under consideration [*Woodhouse and Wong*, 1986; *Romanowicz*, 1987]. Using these results, the propagator terms $\exp i\,(ka\,\Delta+\pi/4)/\sqrt{\sin\Delta}$ in (1) and (5) should for the case of a smooth reference medium be replaced by

$$ka\,\Delta \;\rightarrow\; ka\left[\Delta - \int_0^{\Delta} \frac{\delta c}{c}\, d\Delta'\right] \qquad (10a)$$

$$\sin\Delta \;\rightarrow\; \sin\Delta - \int_0^{\Delta} \sin\Delta' \sin(\Delta-\Delta')\, \partial_{nn}\left(\frac{\delta c}{c}\right) d\Delta' \qquad (10b)$$

The azimuth terms in the polarization vectors and the scattering angle should be replaced by

$$\mu_S \;\rightarrow\; \mu_S - \frac{1}{\sin\Delta} \int_0^{\Delta} \sin(\Delta-\Delta')\, \partial_n\left(\frac{\delta c}{c}\right) d\Delta' \qquad (10c)$$

$$\mu_R \;\rightarrow\; \mu_R + \frac{1}{\sin\Delta} \int_0^{\Delta} \sin\Delta'\, \partial_n\left(\frac{\delta c}{c}\right) d\Delta' \qquad (10d)$$

with similar expressions for the azimuths of the incoming and outgoing wave at the scatterer. In these expressions, $\delta c/c$ is the relative phase velocity perturbation of the reference medium, while ∂_n and ∂_{nn} are the first and second angular derivatives in the transverse direction.

One should be careful giving $\mathbf{u}^1$ the interpretation of the scattered surface wave because $\mathbf{u}^1$ describes all perturbations of the wave field due to the perturbations superposed on the reference medium. If there are abrupt lateral variations, this leads to surface wave scattering. However, in the case of a smoother perturbation on the reference model, $\mathbf{u}^1$ describes the change in the direct wave due to these inhomogeneities. For example, it is shown explicitly by *Snieder* [1987*b*] that the "scattering integral"

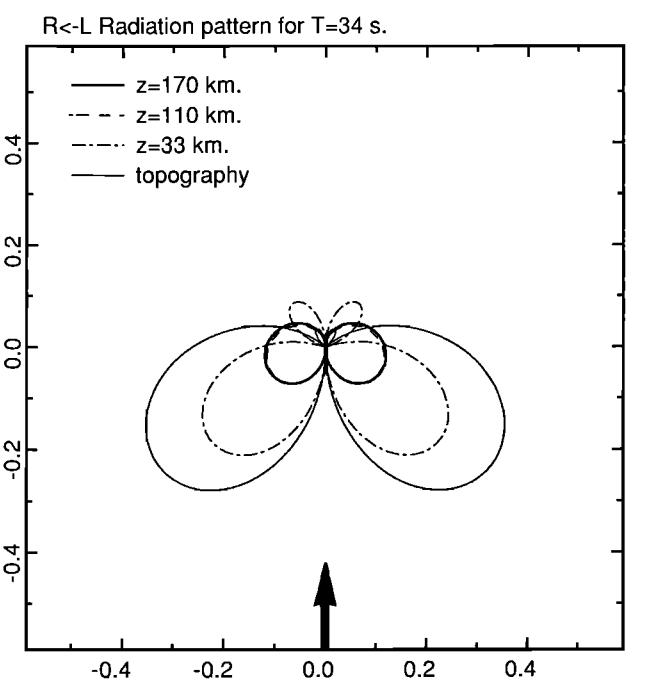

Fig. 1*b*. Radiation pattern for the conversion from the fundamental Love mode to the fundamental Rayleigh mode. Conventions as in Figure 1*a*.

(5) describes the ray geometrical effects on the direct wave due to smooth lateral heterogeneity.

3. ISOTROPIC APPROXIMATION

The surface wave scattering formalism, as presented in section 2, establishes a linear relation between the lateral heterogeneity and the perturbations of the surface wave field. In principle, a three-dimensional inversion could therefore be formulated as a huge system of linear equations by discretizing both the scattering integral over the heterogeneity (5) and the depth integrals in the interaction terms (7). Unfortunately, the simplicity of this approach is elusive. An inversion using the surface wave scattering integral (5) should take care of the following effects: (1) The inhomogeneities should be located at their correct horizontal position. (2) The depth distribution of the heterogeneity should be determined. (3) The contributions from the different inhomogeneities $\delta\rho$, $\delta\lambda$, and $\delta\mu$ should be unravelled. It is difficult those achieve these goals, since the heterogeneity acts on the wave field only through the interaction terms $\tilde{V}^{\nu\sigma}$. This means that it is only possible to retrieve certain depth integrals of the heterogeneity. Information for different frequencies, and possibly different modes, is needed for the reconstruction of the depth dependence of the inhomogeneity. The contribution of the different types of inhomogeneity ($\delta\rho$,$\delta\lambda$,$\delta\mu$) can only be retrieved by using information of different scattering angles.

It will be clear that a complete three-dimensional reconstruction of the heterogeneity is hard to realize with a finite set of band-limited, noise-contaminated data. With present data sets there are two realistic approaches. One can parameterize the depth dependence and the different contributions of $\delta\rho$, $\delta\lambda$, and $\delta\mu$ in a finite set of basis functions. This reduces the degrees of freedom of the heterogeneity, which facilitates a well-behaved inversion. This approach has been taken in a field experiment where surface waves on a tidal flat were reflected by a concrete dam [*Snieder*, 1987*a*]. In this test example the depth dependence of the heterogeneity was prescribed, and an accurate reconstruction of the location of the dam was realized using the surface wave coda.

Alternatively, one can make the "isotropic approximation." It follows from (8) that the $R \leftarrow R$ radiation pattern is stationary with respect to the scattering angle for near forward directions. This can be verified in Figure 1*a* for several different inhomogeneities. Furthermore, it follows from (7) that the $R \leftarrow L$ conversion vanishes in the forward direction. From Figure 1*b* it can be seen that for the shown examples the $R \leftarrow L$ conversion is small for near forward directions. This means that (at least for the fundamental modes) for near forward directions one can make the "isotropic approximation." This means that

$$\tilde{V}_{RL} \approx 0 \tag{11}$$

$$\tilde{V}_{RR} \approx \tilde{V}_{RR}^{(0)} + \tilde{V}_{RR}^{(1)} + \tilde{V}_{RR}^{(2)} = -\left[\frac{l+\frac{1}{2}}{2\pi}\right]^{\frac{1}{2}} (l+\tfrac{1}{2}) \frac{\delta c}{c} \tag{12}$$

These expressions are extremely useful because they make it possible to retrieve the phase velocity perturbation from scattering theory. This allows a two-stage inversion of surface wave data. In the first step the scattering theory is used to find the phase velocity perturbation using (5), (11), and (12). Once these local phase velocities are computed, a standard linear inversion can be used to determine the depth dependence of the heterogeneity [*Nolet*, 1981]. The catch is that this approach forces us to use information for small scattering angles only. In practice, this can be achieved by time windowing the seismograms, and only using the information contained in the direct wave. Note that there are no smoothness restrictions on the heterogeneity, so that it is in principle possible to reconstruct a two-dimensional phase velocity field without doing any dispersion measurements. In this way, the conditions for the validity of the great circle theorem need not be fulfilled.

4. INVERSION OF THE SCATTERING INTEGRAL

The linear relation (5) between the perturbation of the wave field and the perturbation of the medium can be written as

$$\mathbf{u}^1 = \sum_{\nu,\sigma} \iint \mathrm{p}^{\nu} g_{\nu}(\Delta_2) \frac{\partial \tilde{V}^{\nu\sigma}}{\partial m} \cdot m(\theta',\phi') \, g_{\sigma}(\Delta_1) \, (\mathbf{E}^{\sigma *} : \mathbf{M}) \, d\Omega' \tag{13}$$

with the propagators defined by

$$g_{\nu}(\Delta) = \frac{e^{i(k_{\nu} a \Delta + \frac{\pi}{4})}}{\sqrt{\sin \Delta}} \tag{14}$$

or its equivalent for a smoothly varying reference medium (10). The model parameter m designates either the heterogeneity ($\delta\rho$,$\delta\lambda$,$\delta\mu$) parameterized in some suitable form or the phase velocity perturbation $\delta c/c$ if the isotropic approximation is used. The difference between the recorded surface wave data and the synthetics for the reference model ($\mathbf{u}^0$) can for all events, stations, and frequency components be arranged in one (huge) vector $\mathbf{d}$ of data residuals. Likewise, the model parameters can, after a discretization in cells of the surface integral (13) (and possibly also of the depth integrals in the interaction terms), be arranged in one model vector $\mathbf{m}$. (Of course, one does not have to expand the heterogeneity in cells; other parameterizations can also be used.) In that case, (13) can be written as a matrix equation

$$d_i = \sum_j G_{ij} m_j \tag{15}$$

where G_{ij} is the spectral component of the synthetic seismogram for event-station pair "i" at frequency ω_i, due to a unit perturbation of model parameter "j."

In general, the matrix $\mathbf{G}$ is extremely large. The reason for this is that the integrand in the original scattering equation (13) is rapidly oscillating with the position of the inhomogeneity. This means that in order to discretize (13) accurately, a cell size much smaller than a wavelength is needed. For an inversion on a continental scale for surface waves with a wavelength of say 100 km, several thousands of cells are needed. Fortunately, extremely large systems of linear equations can be solved iteratively in the least squares sense (Van der Sluis and Van der Vorst, 1987), so that a brute force inversion of $\mathbf{G}$ need not be performed. The least squares solution minimizes the misfit $|\mathbf{d}-\mathbf{G}\mathbf{m}|^2$, so that one performs in fact a least squares waveform fit of the data residual $\mathbf{d}$ to the synthetics $\mathbf{G}\mathbf{m}$.

In the inversions presented in this paper, and in paper 2, the algorithm LSQR of *Paige and Saunders* [1982*a*,*b*] is used to solve (15) iteratively in the least squares sense. LSQR performs the inversion by doing suitable matrix multiplications with $\mathbf{G}$ and $\mathbf{G}^T$. (In the language of modern optimization schemes [*Tarantola and Valette*, 1982], one would say that one only needs to solve the forward problem.) There is no need to store the matrix in memory; in fact, one only needs to supply LSQR with a subroutine to do a multiplication with one row of $\mathbf{G}$ or $\mathbf{G}^T$. As an additional advantage, LSQR has convenient "built in" regularization properties [*Van der Sluis and Van der Vorst*, 1987]. The stability of LSQR is confirmed by *Spakman and Nolet* [1987], who applied LSQR to a tomographic inversion of an extremely large set of P

wave delay times and who made a comparison with other iterative solvers of linear equations.

The inversion with LSQR has some interesting similarities with migration methods in exploration seismics. The first iteration of LSQR yields a solution proportional to $\mathbf{G}^T\mathbf{d}$, higher iterations perform corrections to the misfit [*Van der Sluis and Van der Vorst*, 1987]. It is shown in detail by *Snieder* [1987*a*] that the contraction $\mathbf{G}^T\mathbf{d}$ amounts to a holographic reconstruction of the heterogeneity. This means that the waves propagating away from the sources (the illumination) are correlated with the surface wave residuals which have back propagated from the receivers into the medium. For one source-receiver pair this leads to an ellipsoidal contribution to the reconstructed image. By summing over all source-receiver pairs (which is implicit in the product $\mathbf{G}^T\mathbf{d}$) an image is constructed. It is shown by *Tarantola* [1984*a*,*b*] that this procedure is similar to Kirchhoff migration as used in exploration seismics. Just as with these techniques, the surface wave reconstructions using the method of this paper will contain "smiles" [*Berkhout*, 1984] if insufficient data are used.

It may be advantageous to impose an a priori smoothness constraint on the solution. This can be achieved by solving instead of (15) the matrix equation

$$\mathbf{GS}\tilde{\mathbf{m}} = \mathbf{d} \tag{16}$$

where $\mathbf{S}$ is a prescribed smoothing matrix. This yields the solution

$$\mathbf{m} = \mathbf{S}\tilde{\mathbf{m}} \tag{17}$$

which incorporates the smoothness criterion imposed by $\mathbf{S}$.

5. PRACTICAL IMPLEMENTATION OF SOLVING THE MATRIX EQUATION

Solving the linear system (15) or (16) is not entirely straightforward because the matrix may be extremely large. For example, discretizing the continent of Europe (with a size of say 3500×3500 km^2) in cells of 35×35 km^2 (which is 1/4 of the wavelength of a 30-s fundamental mode Rayleigh wave) leads to a model of 10,000 cells. For the data set used in paper 2, there are approximately 2500 spectral components of surface wave data to be fitted. This means that storing this matrix requires 100 Mbyte of disc space, which is impractical (if not impossible on many machines). As mentioned before, LSQR does not need the whole matrix at once but only needs access to the rows of $\mathbf{G}$ and $\mathbf{G}^T$. In principle, the matrix can therefore be computed during the inversion. However, due to the large number of trigonometric operations required for the computation of the synthetics this leads to prohibitive CPU times.

If we restrict ourselves to vertical component data for the fundamental mode only, the elements of the matrix $\mathbf{G}$ have the form

$$G = A_R e^{i\phi_R} + A_L e^{i\phi_L} \tag{18}$$

The first term in this expression describes the scattering of the fundamental Rayleigh mode to itself, while the second term describes the conversion from the fundamental Love mode to the fundamental Rayleigh mode. The terms ϕ_R and ϕ_L are the phase terms of the propagators (14), while the complex amplitudes A_R and A_L contain the remaining terms.

Due to the phase terms, the matrix G_{ij} (which is the synthetic for data point i due to a unit perturbation of model parameter j) is an oscillatory function of the position of the inhomogeneity and hence of the index j. This oscillatory character makes it impossible to use some interpolation scheme to compute G_{ij}. However, the phase functions ϕ_R, ϕ_L and the complex amplitudes A_R, A_L are smooth functions of the location of the inhomogeneity. This makes it possible to store these terms at selected grid points and to compute values at intermediate points by interpolation during the inversion. One could call this procedure "Filon matrix multiplication."

This procedure can be simplified even further by using the fact that the the wave numbers of the fundamental Rayleigh wave and the fundamental Love wave usually are not too different (hence $\phi_R \approx \phi_L$). If (18) is written as

$$G = Z\, e^{i\phi_R} \tag{19}$$

with

$$Z = A_R + A_L\, e^{i(\phi_L - \phi_R)} \tag{20}$$

one only needs to store Z and ϕ_R at selected grid points.

The functions Z and ϕ_R are, in general, also a smooth function of frequency, so that the matrix only needs to be stored at certain selected frequencies. The value of the matrix elements for intermediate frequencies can also be computed by interpolation. This interpolation with respect to frequency can be performed with a simple linear interpolation. For the interpolation with respect to the location of the inhomogeneity it is better to use a quadratic scheme. The reason for this is that the phase ϕ_R has a minimum on the minor arc between the source and the receiver. It is especially at this location that accuracy is required if the isotropic approximation is used, because in that case the requirement of a small scattering angle confines the solution to the vicinity of the minor arc. A linear interpolation scheme for the horizontal coordinates is not able to reproduce such a minimum and is thus unsuitable.

In the inversions shown in paper 2 for the structure under Europe and the Mediterranean, an area of 3500×3500 km^2 is investigated. Storing the matrix on a 15×15 grid and interpolating in between produced accurate results. (Halving the grid distance for the interpolation did not change the solution.) In the period range from 30 s to 100 s, only 15 frequencies were sufficient to achieve an accurate interpolation with respect to frequency. (Doubling the number of frequency points for the interpolation did not change the solutions.) In this way, only 2 Mbyte of disc space was needed to store the interpolation coefficients for the matrix $\mathbf{G}$.

The edges of the domain of inversion require special attention. Artificial reflections may be generated at the edge of the domain of inversion if this domain is truncated abruptly (E. Wielandt, personal communication, 1986). This problem can be circumvented by tapering the matrix $\mathbf{G}$ near the edges of the domain. In the inversions used in this study a linear taper was applied to $\mathbf{G}$ near the edges of the domain over a length of 254 km.

The theory formulated here is strictly valid only in the far field [*Snieder and Nolet*, 1987]. It can be seen from (14) that the theory becomes singular in the near field. Due to the lack of a better theory, this problem is ignored in this study. The singularity was removed by replacing the $\sin\Delta$ term in the propagator (14) by a constant ($\sin\Delta_o$), whenever $\Delta<\Delta_o$. A value of 2.7° was adopted for Δ_o.

The data fit (15) or (16) is performed in the frequency domain, whereas surface wave data are recorded in the time domain. After applying some taper these data can be transformed to the frequency domain. In case one uses the isotropic approximation a time window is needed to extract only the direct wave. In general, the data are therefore in the time domain multiplied with some

nonnegative time window $w(t)$. Of course, the matrix elements, which are the spectral components of the inhomogeneity in each cell, should incorporate the effects of this time window. A multiplicative window in the time domain acts as a convolution in the frequency domain, which complicates the inversion. However, it is shown in Appendix A that if the time window $w(t)$ is nonnegative and sufficiently broad, that due to the surface wave character of the signal this filter acts in the frequency domain as a simple multiplication with $w(L/U(\omega))$. In this expression, L is the distance covered by the surface wave, and $U(\omega)$ is the group velocity of the mode under consideration.

6. WAVEFORM FITTING BY NONLINEAR OPTIMIZATION

The theory presented here establishes an inversion scheme in case a linear relation exists between the inhomogeneity and the deviation between the recorded surface waves and the synthetics for the reference model. In practice, this relation may suffer from nonlinearities. The main culprit for this effect is that small changes in the wave number are multiplied in the exponent by a large epicentral distance so that $\exp i(k+\delta k)L \approx (1+i\,\delta kL)\exp ikL$ may be a poor approximation.

It is therefore desirable to perform first a nonlinear inversion in order to find a smooth reference model for the Born inversion. This nonlinear inversion can be achieved by minimizing the penalty function

$$F(m)=\sum_{r,s}\int\left[u^{rs}(t)-s^{rs}(m,t)\right]^2 dt+\gamma\iint|\nabla_1 m|^2\,d\Omega \quad (21)$$

with respect to the model parameters m. In this expression, $u^{rs}(t)$ is the surface wave seismogram for source s and receiver r, while $s^{rs}(t)$ is the corresponding synthetic for model m. The last term serves to select the smoothest possible solution by minimizing the horizontal gradient $|\nabla_1 m|$.

As shown by *Nolet et al.* [1986*a*], the minimization of F in (21) can be achieved efficiently using conjugate gradients. In this kind of inversion one only needs to solve the forward problem repeatedly [*Nolet et al.*, 1986*a*], and most of the computer time is spend computing the gradient of the penalty function with respect to the model parameters. It is therefore crucial to have a fast method for computing this gradient. (In this study, the forward problem is solved using the line integrals (10) in order to incorporate ray geometrical effects. Bicubic splines are useful for representing the lateral phase velocity variations because they ensure continuity of the phase velocity with its first and second derivatives. In this approach, the model parameters m are the phase velocities at some selected grid points.)

The gradient of the misfit $M(m)=\int(u(t)-s(m,t))^2dt$ for one source receiver pair can for band limited data be estimated analytically. It is shown in Appendix B that if the synthetic consists of a sum of modes

$$S(m,\omega)=\sum_\nu S_\nu(m,\omega)=\sum_\nu A_\nu e^{i\phi_\nu} \quad (22)$$

the gradient of the misfit can be approximated by

$$\frac{\partial M}{\partial m}=-2\sum_\nu\left\{\frac{1}{c_\nu^2}\int\frac{\partial c_\nu}{\partial m}\,dx\,B_1^\nu+\frac{1}{A_\nu}\frac{\partial A_\nu}{\partial m}B_2^\nu\right\} \quad (23)$$

In this expression

$$B_1^\nu=\int\dot{s}_\nu(m,t)\left[u(t)-s(m,t)\right]dt \quad (24)$$

$$B_2^\nu=\int s_\nu(m,t)\left[u(t)-s(m,t)\right]dt \quad (25)$$

Model for scattering computation.

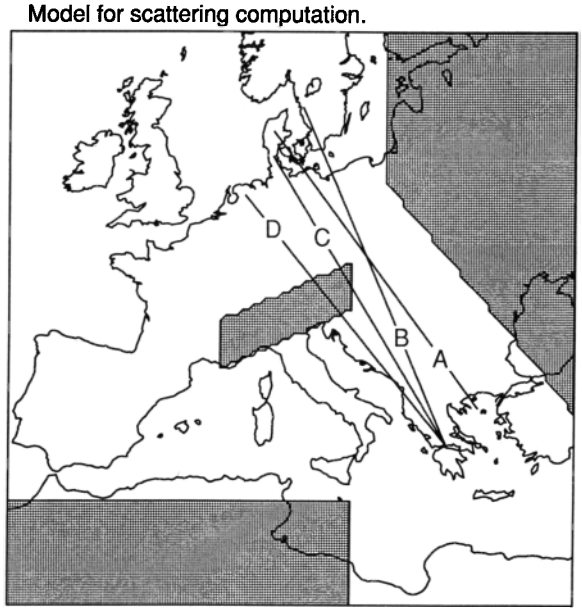

Fig. 2. Horizontal extend of the heterogeneity used in the example of the scattering computation. The inhomogeneity extends down to 170 km with $\delta\beta/\beta=10\%$, $\delta\rho=0$, and $\delta\lambda=\delta\mu$.

and $\int.....dx$ denotes the integral over the minor arc from the source to the receiver. The virtue of this approach is that the correlations B_1^ν and B_2^ν have to be computed only once and that the derivatives of all model parameters follow from these correlations. Note the similarity between (24) and the correlation functions used by *Lerner-Lam and Jordan* [1983] (the "bccfs") in their linear inversion of surface wave data.

7. A NUMERICAL EXAMPLE OF SCATTERED SURFACE WAVES

In order to see whether the scattering theory presented here is useful for inversion, it is instructive to study synthetic seismograms for some artificial distribution of scatterers. Figure 2 shows a fictitious distribution of scatterers which forms an extremely crude model of structures as the Alps, the Tornquist-Teisseyre zone, and the edge of north Africa. As a reference structure, the M7-model of *Nolet* [1977] is used for the density and the elastic parameters, while the attenuation of the PREM model [*Dziewonski and Anderson*, 1981] is employed. The inhomogeneity consists of a constant S velocity perturbation of 10% down to a depth of 170 km, while the density is unperturbed. Equal perturbations of the Lamé parameters are assumed. Synthetics are computed with a brute force integration of (5). In order to satisfy the criterion of linearity, only periods larger than 30 s are considered (see paper 2).

Figure 3 shows the synthetic seismogram for the laterally homogeneous reference model, the model with the inhomogeneity, and data recorded at station NE02 of the NARS network [*Dost et al.*, 1984; *Nolet et al.*, 1986*b*]. Observe the realistic looking coda in the synthetics for the model with the scatterers. Of course, one cannot speak of a fit of the recorded surface wave data for this simple minded model, but the coda in the data and in the synthetics are at least of the same nature. Given the group velocity of the surface waves, the contributions from the different scatterers can be identified by the arrival time of the surface waves. The surface waves scattered by the "Tornquist-Tesseyre zone" interfere with the later part of the

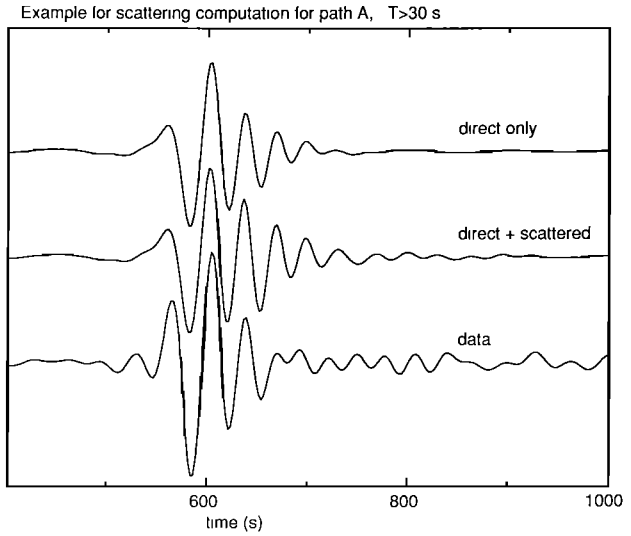

Fig. 3. Seismograms for path A of Figure 2. The top seismogram is for the laterally homogeneous reference medium, the middle seismogram is for the medium with the heterogeneity, and the bottom seismogram shows data.

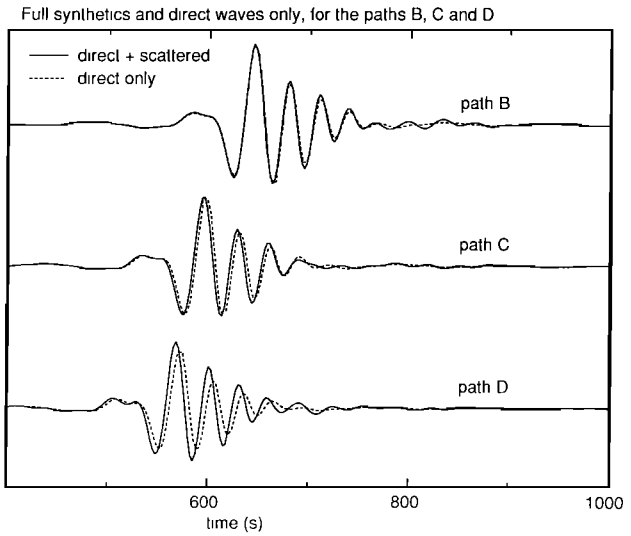

Fig. 4. Synthetic seismograms for the laterally homogeneous reference medium and the medium with the inhomogeneity for the paths B, C, and D of Figure 2.

direct wave and lead to an amplitude increase of approximately 40% of the direct surface wave train around 650 s. The surface waves scattered by the western side of the Alpine block and the diffraction by the corner of the African block constitute the surface wave coda between 750 and 950 s. The diffraction by the corner of the African block (arriving around 900 s) is rather weak because the interaction terms for the corresponding scattering angle are relatively small (see Figures 1*a* and 1*b*). However, the surface wave coda can be made arbitrarily strong by varying the strength and the location of the scatterers and by allowing shorter perids to contribute.

In Figure 4 the synthetics are shown for paths which propagate with different lengths through the central block which mimics the Alps. For path B, which does not propagate through the heterogeneity, only the coda is affected, while for the paths C and D the direct wave is substantially distorted. For path D the inhomogeneity induces both a forward time shift as well as an amplitude increase. Physically, this happens because the scattered waves arrive almost simultaneously with the direct wave (forward scattering). The resulting interference leads to a distortion of the arriving wave train. This example shows that nonsmooth structures may lead to a distortion of the direct surface wave. Interestingly, the phase shift of the direct surface wave in seismogram D coincides up to a deviation of approximately 15% with the path-averaged value of the phase velocity perturbation. This implies that in this case the phase of the direct surface wave is described well by ray theory, despite the fact that applying ray theory is strictly not justified. However, the amplitude of the surface wave is very sensitive to abrupt lateral variations of the structure.

8. INVERSION FOR A POINT SCATTERER

In order to see how the inversion for the surface wave coda operates, an example is shown where one point scatterer influences one seismogram. This point scatterer has the same depth structure as in the example of section 7, but has an effective strength of $\delta\beta/\beta \times \text{area} = 70 \times 70\ \text{km}^2$. The synthetics for the laterally homogeneous reference medium and the (synthetic) data for the medium with the scatterer are shown in the top seismograms of Figure 5. The point scatterer has generated a wave packet which arrives after the direct wave between 600 and 700 s. The Born inversion is applied to these data for a model of 100×100 cells. After three iterations the model shown in Figure 6 is produced. (The correct depth dependence of the heterogeneity is prescribed.) The corresponding synthetics are shown in the bottom seismograms of Figure 5. The "data" for this point scatterer have been fitted quite well.

The resulting model (Figure 6) bears, of course, no resemblance to the original point scatterer because it consists of an ellipsoidal band of positive and negative anomalies. With one source and one receiver it is impossible to determine the true location of the heterogeneity on this ellipse. By using more sources and

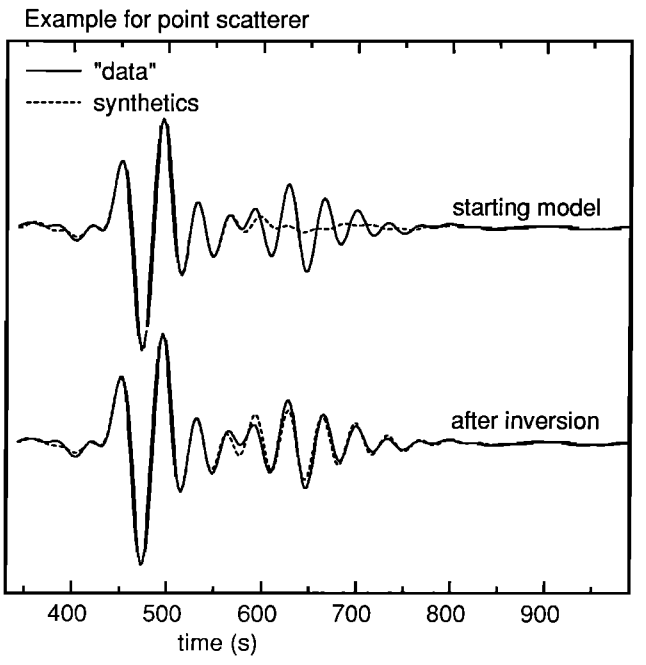

Fig. 5. Waveform fit before and after Born inversion for a synthetic seismogram generated with the point scatterer of Figure 6.

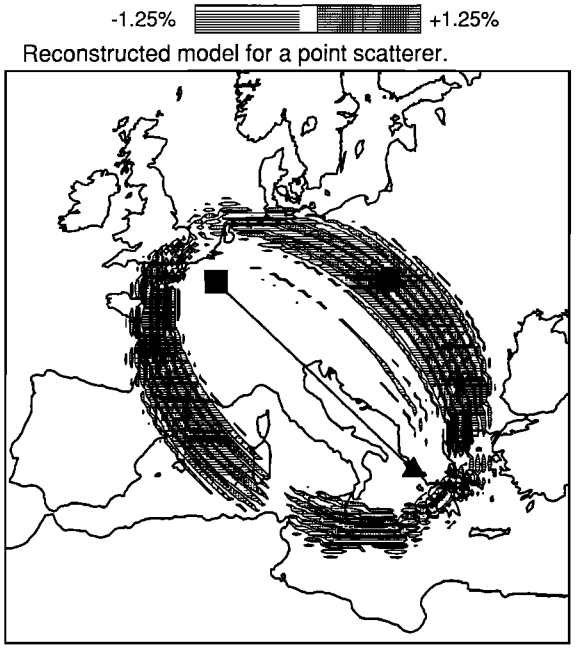

Fig. 6. Relative shear velocity perturbation ($\delta\beta/\beta$) as determined from Born inversion of the top (solid) seismogram of Figure 5. The triangle marks the source, the square marks the receiver, and the circle gives the true location of the point scatterer.

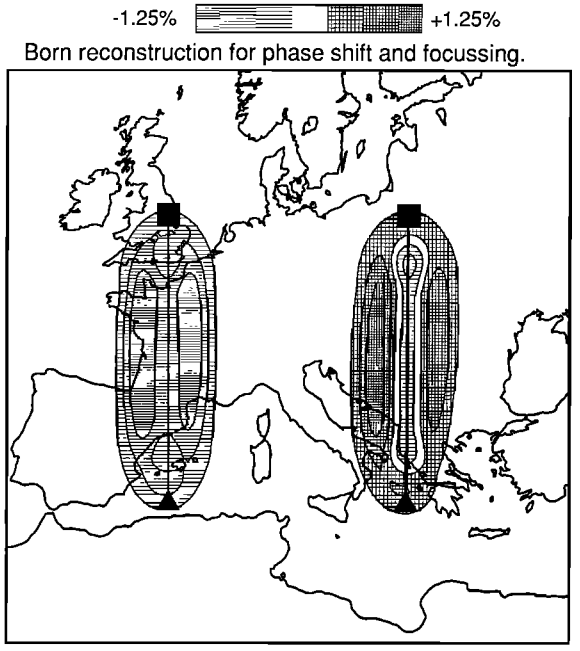

Fig. 7. Relative phase velocity perturbation ($\delta c/c$) as determined from Born inversion for ray geometrical effects. The triangles indicate the sources; the squares mark the receivers.

receivers, an image is constructed by the superposition of these ellipses.

As mentioned in section 4, the result of the first iteration of the Born inversion is proportional to $\mathbf{G}^T\mathbf{d}$, which can be interpreted as the temporal correlation between the excited wave field and the back propagated data residuals [*Snieder*, 1987*a*]. Since surface wave trains consist of oscillating wave packets, this correlation also has an oscillatory nature, which produces the alternation of positive and negative anomalies in Figure 6. The "holes" in these ellipses are caused by the nodes in the radiation pattern of the source (a double couple) and in the radiation pattern of the scatterers (the thick solid curve in the Figures 1*a* and 1*b*).

The strength of the reconstructed inhomogeneity is of the order of 1%, whereas the synthetic "data" have been computed for a point inhomogeneity of 100% with an effective area of 70×70 km^2. The reconstructed heterogeneity is spread out over a much larger area, which explains the weakened reconstructed image. Suppose the heterogeneity is spread out over zone of 2000×300km^2, which is about the right size (see Figure 6). This would lead to a weakening of the reconstructed image of 70×70 km^2/2000×300 km$^2 \approx 1\%$, which is of the order of magnitude of the reconstruction in Figure 6.

9. INVERSION FOR RAY GEOMETRICAL EFFECTS

In this section it is shown how the Born inversion takes ray geometrical effects such as focusing and phase shifting into account. Synthetics have been computed for the two source-receiver pairs shown in Figure 7, assuming a double-couple source for the excitation. The seismogram for the right wave path has been multiplied with 1.4, and the seismogram for the left wave path has been shifted backward in time over 4 s (which is roughly 1% of the travel time).

These seismograms have been inverted simultaneously with the Born inversion using the isotropic approximation. In this inversion a smoothness criterion is imposed because, in contrast to the scattering example of section 8, no sharp heterogeneities are needed to generate the perturbations of the wave field. The domain shown in Figure 7 consists of 100×100 cells of 35×35 km^2. The smoothing matrix that is used is given by

$$S_{i_\phi i_\theta, j_\phi j_\theta} = \alpha^{|i_\phi - j_\phi|}\,\alpha^{|i_\theta - j_\theta|} \quad \text{if} \quad |i_\phi - j_\phi| \le N \quad \text{and} \quad |i_\theta - j_\theta| \le N$$
$$S_{i_\phi i_\theta, j_\phi j_\theta} = 0 \quad \text{elsewhere} \qquad (26)$$

where i_ϕ, i_θ, etc., denote the cell indices in the horizontal directions. In this example the values α=0.66 and N=4 are adopted.

The resulting model after three iterations is shown in Figure 7. Note that because the isotropic approximation is used, Figure 7 displays the phase velocity perturbation $\delta c/c$. (In the inversion a constant value of $\delta c/c$ over the whole frequency band is assumed.) In Figure 8 the waveform fit for the left wave path in

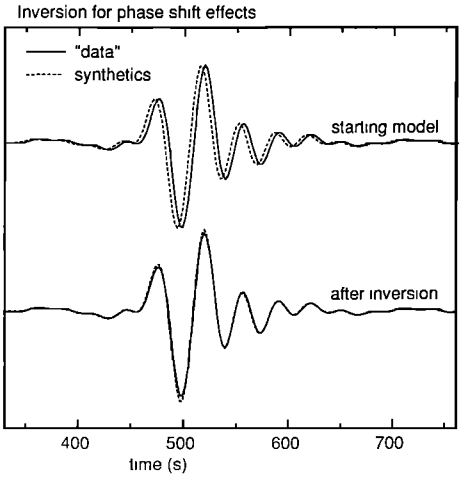

Fig. 8. Waveform fit before and after Born inversion for the left wave path in Figure 7, where the phase is shifted.

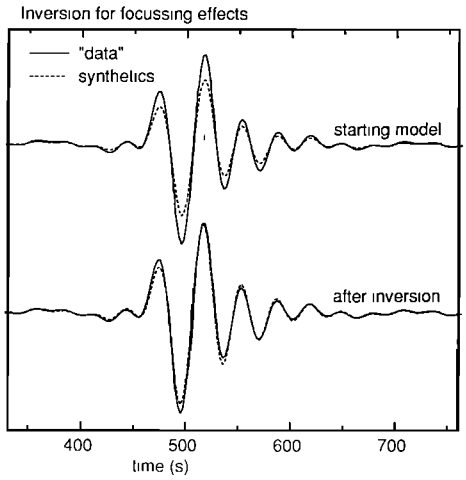

Fig. 9. Waveform fit before and after Born inversion for the right wave path in Figure 7, where the amplitude is increased.

shown. (This is the time shifted seismogram.) The phase shift is correctly taken into account. This is realized by a negative phase velocity anomaly in the first Fresnel zone of the left wave path in Figure 7. This negative phase velocity anomaly is not distributed evenly over the first Fresnel zone of the left wave path; there are phase velocity minima slightly away from the source receiver minor arc. If these minima were absent, the resulting concave transverse phase velocity profile would produce an anomalously large amplitude due to focusing. Because of the phase velocity minima adjacent to the source receiver minor arc, the transverse phase velocity profile is actually convex at the minor arc. This produces defocusing of surface wave energy, which compensates the amplitude increase due to refraction at the edge of the Fresnel zone.

The seismograms for the right wave path are shown in Figure 9. The synthetic data are 40% too strong for the laterally homogeneous reference model; this is almost completely taken care of in the inversion. Physically, this is achieved by a negative phase velocity anomaly on the source receiver line and an anomalously high phase velocity just away from this line. This phase velocity pattern leads to focusing of surface wave energy, so that the large amplitude is fitted. This confirms not only that surface wave scattering theory can account for ray geometrical effects [*Snieder*, 1987*b*], but also that these ray geometrical effects are taken care of in the Born inversion. The asymmetry of the phase velocity pattern in Figure 7 around the wave paths is due to the the asymmetry in the radiation pattern of the double-couple source.

There are approximately 10 cells between the maxima in the strips of high phase velocities for the right wave path in Figure 7. The focusing produced by this structure is achieved by the transverse curvature of the phase velocity. Increasing the cell size (which is computationally advantageous) leads to a representation of this curvature with only a few cells, which may produce unacceptable inaccuracies.

10. CONCLUSION

Large-scale inversion of the surface wave coda can in principle be performed using an iterative solver of a large system of linear equations. For this kind of inversion the depth dependence of the heterogeneity should be prescribed or be parameterized in a limited number of basis functions. Alternatively, the isotropic approximation can be used, which leads to a waveform fit of the direct surface wave due to a laterally heterogeneous phase velocity field. These phase velocities, determined for different frequency bands, can be inverted locally to a depth distribution of the heterogeneity.

The Born inversions shown in this paper are performed iteratively using LSQR. Although LSQR is originally designed for sparse matrices and the matrix for surface wave scattering is not sparse, good results are obtained in inversions of synthetic data. In practice, three iterations proved to be sufficient both for an inversion for the surface wave coda and of the direct wave. A similar conclusion was drawn by *Gauthier et al.* [1986], who used an iterative scheme for fitting waveforms in an exploration geophysics setting.

In hindsight, the success of linear waveform inversions in a few number of iterations is not so surprising. It has been argued by *Tarantola* [1984*a*,*b*] that the standard Kirchhoff migration methods in exploration seismics is equivalent to the first (steepest descent) step of an iterative optimization scheme. Analogously, the first step of the iterative matrix solver used here amounts to a holographic inversion [*Snieder*, 1987*a*] analogously to Kirchhoff migration. These one-step migration methods have been extremely successful in oil exploration, and there is no principal reason why a similar scheme cannot be used in global seismology. Applications of this technique to surface wave data recorded by the NARS array are shown in paper 2 [*Snieder*, this issue].

APPENDIX A: EFFECT OF THE TIME WINDOW FUNCTION ON THE SPECTRUM OF SURFACE WAVES

Suppose that a surface wave seismogram $s(t)$ is multiplied with some nonnegative window function $w(t)$ to give a windowed seismogram $f(t)$

$$f(t) = w(t)s(t) \tag{A1}$$

In the frequency domain the application of this window leads to a convolution

$$F(\omega) = \int W(\omega')S(\omega-\omega')\,d\omega' \tag{A2}$$

Since $w(t)$ is nonnegative, $|W(\omega)|$ attains its maximum for $\omega=0$; this can be seen by making the following estimates:

$$|W(\omega)| = |\int w(t)e^{i\omega t}dt| \le \int |w(t)e^{i\omega t}|dt$$
$$= \int w(t)dt = |W(\omega=0)| \tag{A3}$$

If the time window has a length T in the time domain, its frequency spectrum will have a width of the order π/T in the frequency domain. From this we conclude that long nonnegative time windows have a spectrum that peaks around $\omega=0$.

Now assume that the surface wave spectrum consists of one mode (extensions to multimode signals are straightforward):

$$S(\omega) = A(\omega)e^{ik(\omega)L} \tag{A4}$$

where L is the epicentral distance. Substituting in (A2) gives

$$F(\omega) = \int W(\omega')A(\omega-\omega')e^{ik(\omega-\omega')L}\,d\omega' \tag{A5}$$

$W(\omega)$ is a function peaked around $\omega=0$, so that the main contribution to the ω' integral comes from the point $\omega'=0$. Usually, the complex amplitude $A(\omega)$ is a smooth function of frequency, so that one can approximate for small ω'

$$A(\omega-\omega') \approx A(\omega) \tag{A6}$$

The phase term can be analyzed with a simple Taylor expansion

$$k(\omega-\omega')L \approx k(\omega)L - \frac{L\omega'}{U(\omega)} \tag{A7}$$

where $U(\omega)$ is the group velocity of the surface wave mode. Inserting (A6) and (A7) in (A5) gives

$$F(\omega) \approx A(\omega)e^{ik(\omega)L} \int W(\omega')e^{-iL\omega'/U(\omega)}\, d\omega' \tag{A8}$$

With (A4) and the definition of the Fourier transform this leads to

$$F(\omega) \approx S(\omega)\, w(L/U(\omega)) \tag{A9}$$

APPENDIX B: ANALYTICAL ESTIMATION OF THE GRADIENT $\partial M/\partial m$

The misfit between the data $d(t)$ and the surface wave synthetics $s(m,t)$ for model m is in the L_2 norm defined by

$$M \equiv \int |d(t) - s(m,t)|^2 dt \tag{B1}$$

Using Parseval's theorem [*Butkov*, 1968], the misfit has the same form in the frequency domain

$$M = \int |D(\omega) - S(m,\omega)|^2 d\omega \tag{B2}$$

In general, the model m consists of many parameters. The derivative of the misfit with respect to one of these parameters is

$$\frac{\partial M}{\partial m} = -2\,\mathrm{Re}\left\{\int \frac{\partial S(m,\omega)}{\partial m}(D^*(\omega)-S^*(m,\omega))\, d\omega\right\} \tag{B3}$$

Let the surface wave seismogram be given by a superposition of modes ν with complex amplitude A_ν and phase ϕ_ν

$$S(m,\omega) = \sum_\nu S_\nu(m,\omega) = \sum_\nu A_\nu(m,\omega)e^{i\phi_\nu(m,\omega)} \tag{B4}$$

so that

$$\frac{\partial S}{\partial m} = \sum_\nu \left[\frac{1}{A_\nu}\frac{\partial A_\nu}{\partial m} + i\frac{\partial \phi_\nu}{\partial m}\right] S_\nu \tag{B5}$$

According to equation (10), the phase of the surface waves is in a laterally heterogeneous medium

$$\phi(\omega) = \int_0^L k(\omega,x)\, dx \tag{B6}$$

where $k(\omega,x)$ is the local wave number. Differentiation with respect to the model parameter m gives

$$\frac{\partial \phi}{\partial m} = \int_0^L \frac{\partial k}{\partial m}\, dx = -\frac{\omega}{c^2}\int_0^L \frac{\partial c}{\partial m}\, dx \tag{B7}$$

Inserting this in (B5) gives

$$\frac{\partial M}{\partial m} = -2\,\mathrm{Re}\left\{\sum_\nu \int \frac{1}{A_\nu}\frac{\partial A_\nu}{\partial m} S_\nu(m,\omega)\left[D^*(\omega)-S^*(m,\omega)\right] d\omega\right\}$$

$$+2\,\mathrm{Re}\left\{\sum_\nu \int \frac{i\omega}{c_\nu^2}\left[\int_0^L \frac{\partial c_\nu}{\partial m}dx\right] S_\nu(m,\omega)\left[D^*(\omega)-S^*(m,\omega)\right] d\omega\right\} \tag{B8}$$

When one is attempting to find phase velocities by nonlinear optimization, one will usually work with band passed data for which $c(\omega)$ can be assumed to be independent of frequency. In that case, the phase velocity term and the amplitude term can be taken out of the frequency integral. Applying Parseval's theorem once more to the resulting expression gives

$$\frac{\partial M}{\partial m} \approx -2\sum_\nu \frac{1}{c_\nu^2}\left[\int \frac{\partial c_\nu}{\partial m}dx\right]\int \dot{s}_\nu(m,t)\left[d(t)-s(m,t)\right] dt$$

$$-2\sum_\nu \frac{1}{A_\nu}\frac{\partial A_\nu}{\partial m}\int s_\nu(m,t)\left[d(t)-s(m,t)\right] dt \tag{B9}$$

which proves equation (23).

REFERENCES

Backus, G.E., Geographical interpretation of measurements of average phase velocities over great circular and great semi circular paths, *Bull. Seismol. Soc. Am., 54*, 571–610, 1964.

Berkhout, A.J., *Seismic Migration, Imaging of Acoustic Energy by Wave Field Extrapolation, B, Practical Aspects*, Elsevier, Amsterdam, 1984.

Bungum, H., and J. Capon, Coda pattern and multipathing propagation of Rayleigh waves at NORSAR, *Phys. Earth Planet. Inter., 9*, 111–127, 1974.

Butkov, E., *Mathematical Physics*, Addison-Wesley, Reading, Mass., 1968.

Calcagnile, G., and R. Scarpa, Deep structure of the European-Mediterranean area from seismological data, *Tectonophysics, 118*, 93–111, 1985.

Dahlen, F.A., The spectra of unresolved split normal mode multiplets, *Geophys. J. R. Astron. Soc., 58*, 1–33, 1979.

Dost, B., A. van Wettum, and G. Nolet, The NARS array, *Geol. Mijnbouw, 63*, 381–386, 1984.

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

Gauthier, O., J. Virieux, and A. Tatantola, Two-dimensional nonlinear inversion of seismic waveforms: Numerical results, *Geophysics, 51*, 1387–1403, 1986.

Gilbert, F., and A. Dziewonski, An application of normal mode theory to the retrieval of structural parameters and source mechanisms from seismic spectra, *Philos. Trans. R. Soc. London, Ser. A 278*, 187–269, 1975.

Jordan, T.H., A procedure for estimating lateral variations from low frequency eigenspectra data, *Geophys. J. R. Astron. Soc., 52*, 441–455, 1978.

Lerner-Lam, A.L., and T.J. Jordan, Earth structure from fundamental and higher-mode waveform analysis, *Geophys. J. R. Astron. Soc., 75*, 759–797, 1983.

Levshin, A., and K.A. Berteussen, Anomalous propagation of surface waves in the Barentz Sea as inferred from NORSAR recordings, *Geophys. J. R. Astron. Soc., 56*, 97–118, 1979.

Nolet, G., The upper mantle under Western-Europe inferred from the dispersion of Rayleigh wave modes, *J. Geophys., 43*, 265–285, 1977.

Nolet, G., Linearized inversion of (teleseismic) data, in *The Solution of the Inverse Problem in Geophysical Interpretation*, edited by R. Cassinis, pp. 9–38, Plenum, New York, 1981.

Nolet, G., J. van Trier, and R. Huisman, A formalism for nonlinear inversion of seismic surface waves, *Geophys. Res. Lett., 13*, 26–29, 1986*a*.

Nolet, G., B. Dost, and H. Paulssen, Intermediate wavelength seismology and the NARS experiment, *Ann. Geophys., B4*, 305–314, 1986*b*.

Paige, C.G., and M.A. Saunders, LSQR: An algorithm for sparse linear equations and sparse least squares, *ACM Trans. Math. Software, 8*, 43–71, 1982*a*.

Paige, C.G., and M.A. Saunders, LSQR: Sparse linear equations and least squares problems, *ACM Trans. Math. Software, 8*, 195–209, 1982*b*.

Panza, G.F., S. Mueller, and G. Calcagnile, The gross features of the lithosphere-asthenosphere system from seismic surface waves and body waves, *Pure Appl. Geophys., 118*, 1209–1213, 1980.

Romanowicz, B., Multiplet-multiplet coupling due to lateral heterogeneity: Asymptotic effects on the amplitude and frequency of the Earth's normal modes, *Geophys. J. R. Astron. Soc., 90*, 75–100, 1987.

Romanowicz, B., and R. Snieder, A new formalism for the effect of lateral heterogeneity on normal modes and surface waves, II, General anisotropic perturbations, *Geophys. J. R. Astron. Soc., 93*, 91–100, 1988.

Snieder, R., 3D Linearized scattering of surface waves and a formalism for surface wave holography, *Geophys. J. R. Astron. Soc., 84*, 581–605, 1986*a*.

Snieder, R., The influence of topography on the propagation and scattering of surface waves, *Phys. Earth Planet. Inter., 44*, 226–241, 1986*b*.

Snieder, R., Surface wave holography, in *Seismic Tomography, With Applications in Global Seismology and Exploration Geophysics*, edited by G. Nolet, pp. 323–337, D. Reidel, Hingham, Mass., 1987*a*.

Snieder, R., 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 N.J., Vlaar, G., Nolet, M.J.R., Wortel, S.A.P.L. and Cloetingh, pp. 77–83, D. Reidel, Hingham, Mass., 1987*b*.

Snieder, R., Large-scale waveform inversions of surface waves for lateral heterogeneity, 2, Application to surface waves in Europe and the Mediterranean, *J. Geophys. Res., this issue*, 1988.

Snieder, R., and G. Nolet, Linearized scattering of surface waves on a spherical Earth, *J. Geophys., 61*, 55–63, 1987.

Snieder, R., and B. Romanowicz, A new formalism for the effect of lateral heterogeneity on normal modes and surface waves, I, Isotropic perturbations, perturbations of interfaces and gravitational perturbations, *Geophys. J. R. Astron. Soc., 92*, 207–222, 1988.

Spakman, W., and G. Nolet, Imaging algorithms, accuracy and resolution in delay time tomography, in *Mathematical Geophysics, A Survey of Recent Developments in Seismology and Geodynamics*, edited by N.J., Vlaar, G., Nolet, M.J.R., Wortel, S.A.P.L. and Cloetingh, pp. 155–187, D. Reidel, Hingham, Mass., 1987.

Tanimoto, T., The three-dimensional shear wave velocity structure in the mantle by overtone waveform inversion, I, Radial seismogram inversion, *Geophys. J. R. Astron. Soc., 89*, 713–740, 1987.

Tarantola, A., Inversion of seismic reflection data in the acoustic approximation, *Geophysics, 49*, 1259–1266, 1984*a*.

Tarantola, A., Linearized inversion of seismic reflection data, *Geophys. Prospect., 32*, 998–1015, 1984*b*.

Tarantola, A., and B. Valette, Generalized nonlinear inverse problems solved using the least squares criterion, *Rev. Geophys., 20*, 219–232, 1982.

Van der Sluis, A., and H.A. Van der Vorst, Numerical solution of large, sparse linear algebraic systems arising from tomographic problems, in *Seismic Tomography, With Applications in Global Seismology and Exploration Geophysics*, edited by G. Nolet, pp. 49–83, D. Reidel, Hingam, Mass., 1987.

Woodhouse, J.H., and Y.K. Wong, Amplitude, phase and path anomalies of mantle waves, *Geophys. J. R. Astron. Soc., 87*, 753–774, 1986.

Yomogida, K., and K. Aki, Amplitude and phase data inversion for phase velocity anomalies in the Pacific Ocean basin, *Geophys. J. R. Astron. Soc., 88*, 161–204, 1987.

R. Snieder, Department of Theoretical Geophysics, University of Utrecht, P.O. Box 80.021, 3508 TA Utrecht, The Netherlands.

(Received October 13, 1987;
revised April 21, 1988;
accepted April 25, 1988.)